REPLICATION-COMPETENT NON-INDUCED PROVIRUSES IN THE LATENT RESERVOIR INCREASE BARRIER TO HIV-1 CURE

by
Ya-Chi Ho

A dissertation submitted to the Johns Hopkins University School of Medicine in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
September 2013

ABSTRACT

Antiretroviral therapy (ART) fails to cure HIV-1 infection because latent proviruses persist in resting $CD4^+$ T cells. T cell activation reverses latency, but >99% of proviruses are not induced to release infectious virus after maximum *in vitro* T cell activation under standard viral outgrowth assay conditions. These non-induced proviruses are generally considered defective but have never been characterized.

Using limiting dilution near-full length nested PCR and Poisson distribution analysis, we characterized 213 clones of non-induced proviruses from eight aviremic patients on suppressive ART. Most (88.3%) of the non-induced proviruses are defective, including 45.5% large internal deletions, 32.4% APOBEC-mediated G→A hypermutations, 6.6% mutations/deletions in the *cis*-acting element, and 3.8% insertions/nonsense mutations.

Strikingly, 11.7% of the non-induced proviruses have intact genomes. Using direct sequencing and *de novo* genome synthesis, we reconstructed six full-length non-induced proviral clones and demonstrated growth kinetics comparable to four reconstructed induced proviruses from the same patients.

Using luciferase assay to measure non-induced proviral LTR activity, we found that non-induced proviruses have intact promoter function unless they are hypermutated. Using limiting dilution bisulfite sequencing, we found that non-induced proviruses have unmethylated promoters. Using inverse PCR, we found that non-induced proviruses are

integrated into active transcription units. We demonstrated that these non-induced proviruses, though not induced after *in vitro* maximum T cell activation, can be reactivated after repeated stimuli. We propose that maximum T cell activation does not lead to maximum non-induced provirus activation. Rather, activation of non-induced proviruses is stochastic.

Thus, it cannot be excluded that non-induced proviruses may become activated *in vivo*. The discovery of replication-competent non-induced proviruses indicates that the size of the latent reservoir, and hence the barrier to cure, may be ~60-fold greater than previously estimated.

Underestimation of intact proviruses by viral outgrowth assays could be reflected in delayed viral rebound after an apparent "cure", and overestimation of latent reservoir size resulting from detection of defective proviruses by PCR assays could result in prolonged, excessive exposure to toxic latency reversing agents. Thus, the molecular analysis of non-induced proviruses contributes in an important way to HIV-1 eradication efforts.

Advisor:
Robert F. Siliciano, M.D., Ph.D.
Professor of Medicine
Johns Hopkins University School of Medicine.

Reader:
Joel L. Pomerantz, Ph.D.
Associate professor of Biological Chemistry
Johns Hopkins University School of Medicine

PREFACE

I would like to express my deepest appreciation to my advisor, Dr. Robert F. Siliciano, for his inspirational guidance and full support to my intellectual interests and career. He envisions the most significant issues in the field and asks important questions. He not only guided me to solve difficult problems, but even more importantly, provided me the freedom to ask different questions and conduct experiments on my own. He works as a perfectionist – creative, enthusiastic, industrious and conscientious and insightful. But he is extremely kind, encouraging and generous to me, even for negative results and inadequate performance. His dedication to science and education makes him an exceptional physician-scientist and great mentor. It is my privilege to learn from him, who led me all the way to this exciting scientific adventure.

I would like to thank my unofficial advisor, Dr. Janet Siliciano, for her effort and advises on scientific discussions as well as family management. I would like to thank my thesis committee members, Dr. Stuart Ray, Dr. Joel Blankson and Dr. Joel Pomerantz. Dr. Stuart Ray is very smart and kind. Each time after I presented my work to Dr. Stuart Ray, he always provided insightful and important suggestions immediately. These suggestions eventually answered the critical questions that other world experts in the field asked. Dr. Joel Blankson is always available for discussions and insightful for my directions. He guided me from the smallest technical problems to career planning. Dr. Joel Pomerantz is a distinguished scientist whose enthusiasm in research is so contagious and influential. I had the privilege to work closely with him during my rotation, which built the foundation of my scientific curiosity and molecular biology skills.

I appreciate all the members the Siliciano lab, who constantly give me vibrant scientific discussions and help in a very pleasant atmosphere. Particularly, Dr. Liang Shan and Seyed Alireza Rabi gave me so many great ideas. They are extremely smart, innovative, hardworking, outstanding but humble. I thank Dr. Hung-Chih Yang, Dr. Sifei Xing, Dr. Maria Salgado, Kai Deng, Robert Buckheit III, Christopher Pohlmeyer and Sarah Laskey for their encouragement and friendship. I keep learning from their diligence, multi-tasking ability and altruism. I thank Dr. Daniel Rosenbloom and Sarah Laskey for their comprehensive mathematical estimation and modeling of the size of the latent reservoir. I thank Ms. Bonita Grant and Ms. Jun Lai for their effort meeting all the needs from me and this big lab. I would like to thank my collaborators, including Dr. Eitan Halper-Stromberg and Dr. Haiping Hao for their help for deep sequencing.

Special thanks go to Dr. Pierre Coulombe, Dr. M. Christine Zink, Ms. Leslie Lichter and Ms. Colleen Graham of the Cellular and Molecular Medicine PhD training program for giving me the chance to study in this prestigious institute.

I would like to thank the Ministry of Education of Taiwan who offered me a one-year fellowship to study in the United States. I also thank the Howard Hughes Medical Institutions International Student Research Fellowship program who supported my thesis project and two years of research.

The completion of this PhD means tremendous support from my family. My parents spent the past five years helping me taking care of my daughters so that I can concentrate on my research. They could have lived a much better and easier life, but instead they chose to stay with us in the States with restricted access to their friends.

I dedicate all my work to my husband, Dr. Yu-Min Chuang. He is the very person who understands and fully supports my immense interest in scientific research. Although we work on different fields, he often provides insightful suggestions and asks key questions regarding my project. His love, kindness and dedication to both our family and research supported me during the most desperate period of my life.

I am grateful to all my friends, near or far, who supported me regarding science and life. I look forward to working as a physician-scientist, using basic research tools to solve clinically important questions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

Despite prolonged antiretroviral therapy (ART), HIV-1 persists as transcriptionally inactive proviruses in resting memory $CD4^+$ T cells (Chun et al., 1997; Finzi et al., 1997; Wong et al., 1997). This latent reservoir has a long half-life, preventing cure by ART alone (Finzi et al., 1997; Siliciano et al., 2003; Strain et al., 2003). In resting $CD4^+$ T cells, the lack of active forms of key cellular transcription factors (Bohnlein et al., 1988; Duh et al., 1989; Ganesh et al., 2003; Kinoshita et al., 1997; Nabel and Baltimore, 1987; West et al., 2001) and of HIV-1 Tat and its cellular cofactors (Cujec et al., 1997; Herrmann and Rice, 1995; Jones and Peterlin, 1994; Kao et al., 1987; Selby and Peterlin, 1990; Tyagi et al., 2010) limits the initiation and elongation, respectively, of viral transcription (Lassen et al., 2004; Williams and Greene, 2007). The latent reservoir may thus be established when activated $CD4^+$ T cells become infected as they revert back to a resting memory state. In addition, DNA methylation and repressive histone modifications may promote silencing of proviruses (Blazkova et al., 2009; Coull et al., 2000; He and Margolis, 2002; Kauder et al., 2009; Van Lint et al., 1996; Verdin et al., 1993; Williams et al., 2006).

A major approach to eradicating HIV-1 involves reversing latency in patients on ART (Deeks, 2012; Richman et al., 2009). Cells harboring induced proviruses could then be lysed by HIV-1-specific cytolytic T lymphocytes (CTL) (Shan et al., 2012), while new rounds of infection are blocked by ART. Clinical trials exploring this strategy have used the histone deacetylase inhibitors valproic acid (Lehrman et al., 2005), vorinostat (Archin et al., 2009; Archin et al., 2012; Contreras et al., 2009) and panobinostat (Katlama et al., 2013).

Accurate measurement of the latent reservoir is essential for evaluating eradication strategies. If the latent reservoir is completely eradicated, ART can be discontinued without rebound viremia. Interruption before complete eradication will likely result in rebound (Davey et al., 1999) and repopulation of the latent reservoir, nullifying the eradication effort.

The standard assay for latent reservoir size is a viral outgrowth assay (viral outgrowth assay (Finzi et al., 1997; Siliciano and Siliciano, 2005), which measures the frequency of resting $CD4^+$ T cells that produce infectious virus after a single round of maximum *in vitro* T cell activation. Limiting dilutions of resting $CD4^+$ T cells are stimulated with the mitogen phytohemagglutinin (PHA), which reverses latency by inducing T cell activation. Released viruses are expanded by addition of $CD4^+$ T lymphoblasts from HIV-1-negative donors. Culture supernatants are examined for exponential viral growth by enzyme-linked immunosorbant assay (ELISA) for HIV-1 p24. With this assay, the mean frequency of latently infected cells in patients on ART is $\sim 1/10^6$ resting $CD4^+$ T cells (Eriksson et al., 2013).

It has been assumed that latent reservoir size can be assessed with agents like PHA that induce uniform activation of $CD4^+$ T cells (Hermankova et al., 2003; Patel et al., 1988). However, the frequency of latently infected cells detected in the viral outgrowth assay is on average 300 fold lower than the frequency of resting $CD4^+$ T cells that harbor proviruses detectable by PCR (Eriksson et al., 2013). Thus at limiting dilution in the viral outgrowth assay, negative wells contain many proviruses, which we designate *non-induced proviruses*. The non-induced proviruses are generally considered defective. However, they have not been molecularly characterized. The magnitude of

the challenge presented by the latent reservoir depends on whether non-induced proviruses can be induced *in vivo*. We present here the first molecular characterization of non-induced proviruses, the results of which provide disturbing new insights into the difficulty of curing HIV-1 infection.

The latent reservoir in resting $CD4^+$ T cells is the major barrier to HIV-1 eradication, and as efforts to cure the infection proceed, accurate measurement of latent reservoir size will be essential. This study provides a molecular basis for understanding measures of the latent reservoir. Through an analysis of proviruses that did not give rise to infectious virus following a single round of T cell activation (*non-induced proviruses*), we have provided a definitive explanation for the large discrepancy between results of PCR and culture assays of latent reservoir size. In addition, the discovery of intact non-induced proviruses indicates that the size of the latent reservoir may be much greater than previously thought. We reconstructed full-length, intact non-induced proviruses from multiple patients, and all showed growth kinetics comparable to induced proviruses from the same patient and a reference isolate. These intact non-induced proviruses are not detected in standard culture assays but may nevertheless prevent cure. Thus the present study provides new insights into the extent of the challenge posed by the latent reservoir and may lead to novel strategies that target intact non-induced proviruses.

Chapter 2. Viral outgrowth assay conditions achieves maximum *in vitro* T cell activation

Introduction

For determining the size of the latent reservoir, the current standard is the viral outgrowth assay (86 Siliciano,J.D. 2005; 220 Finzi,D. 1997). This assay measures the frequency of resting $CD4^+$ T cells that produce replication-competent virus after as single round of maximum *in vitro* T cell activation. Limiting dilutions of resting $CD4^+$ T cells purified from patients on suppressive ART are activated with PHA. PHA, a potent mitogenic lectin, cross-links T cell receptor/CD3 complex, activates phospholipase C and protein kinase C (Mustelin et al., 1990), up-regulates interleukin-2 (IL-2) and IL-2 receptor (CD25) expression (Depper et al., 1985), resulting in an autocrine stimulation and T cell proliferation. PHA provides 100% colony forming ability in $CD4^+$ T cells (Hermankova et al., 2003; Patel et al., 1988) and reverses latency by global T cell activation, as one of the most potent reagent used *in vitro* to reverse latency (Yang et al., 2009b), in a greater efficiency than most latency reversing agents tested to date. PHA was washed out after one day of culture to reduce cytotoxicity. Viruses released into the supernatant following PHA activation are further amplified by co-culture with two additions of allogeneic $CD4^+$ T lymphoblasts. Culture supernants are examined for exponential viral growth p24 ELISA for HIV-1 Gag antigen on day 14 – 21 of culture, and positive wells are considered to have contained at least one latently infected cells harboring inducible, replication competent virus. Poisson statistics are used to establish

the frequency of latently infected cells (Siliciano and Siliciano, 2005). In patients on ART, it is typically on the order of one cell per 10^6 resting $CD4^+$ T cells.

Unlike human endogenous retroviruses, which are silenced during embryogenesis (Rowe et al., 2010) and generally thought to be harmless, HIV-1 integrates into terminally differentiated $CD4^+$ T cells and does not follow silencing mechanisms seen in endogenous retroviruses (Rowe et al., 2010). Although they were not induced by *in vitro* activation using PHA, it does not mean that they would never be reactivated *in vivo*.

Unlike PCR based methods, which solely represent the presence of DNA fragment regardless of being intact or defective, the relatively high detection limit of viruses in the supernatant indicate the presence of multiple rounds of exponential viral propagation, reflecting presence of functional and replication competent proviruses. The detection limit of the viral outgrowth assay is 104,167 HIV-1 RNA copies/mL, which was calculated as: 3 fg HIV-1 p24/50 RNA copies (Barletta et al., 2004) x 6.25 pg/mL detection limit = 104,167 copies/mL. However, with this relatively high detection limit, p24 negative wells are not merely a reflection of false negative results of an insensitive detection assay. Rather, it has been confirmed that all p24 negative wells remain negative after a more sensitive mRNA detection method (Laird et al., 2013) down to 555 copies/mL (which was calculated by the fact that detecting one copy/well of quantitative PCR equals to 1.8 µL of culture supernatant in this assay).

The main obstable to isolate full-length proviral DNA from the viral outgrowth cultures is that the patient resting $CD4^+$ T cells (for example 0.2 or 0.04 million per well) are outnumbered by the healthy donor lymphoblasts (1 million per well x 2 additions = 2 million per well). Isolating proviral DNA from this large amount of human chromosomal

DNA is technically challenging. An alternative way is to separate patient cells and healthy donor lymphoblasts by a transwell culture system, with a 0.4 μm membrane permeable for viruses and nutrients but impermeable for cells, to obtain a pure population of patient cells for further proviral DNA isolation. This separation, however, prevents direct intercellular contact, which may or may not reduce the efficiency of viral propagation, and hence the sensitivity of the viral outgrowth assay. We here demonstrated that our transwell viral outgrowth assay provides maximum T cell activation and comparable sensitivity with the standard viral outgrowth assay.

Methods

Viral outgrowth assay

Viral outgrowth assays were performed as described (Finzi et al., 1997; Siliciano and Siliciano, 2005). Briefly, resting $CD4^+$ T cells were isolated from patients under suppressive ART with undetectable plasma viral load for >6 months prior to enrollment to ensure no preintegration latency. Using magnetic microbead selection (Miltenyi Biotec), $CD4^+$ T cells were first isolated using negative selection, and CD25, CD69 and HLA-DR expressing cells were removed by positive selection. Standard and transwell cultures were preformed side by side using the same population of patient resting $CD4^+$ T cells, irradiated allogeneic healthy donor peripheral blood mononuclear cells (PBMCs) and CD8-deplted healthy donor lymphoblasts. The volume of culture medium and the total number of cells per well was equal for these two culture conditions. In transwell cultures, patient resting $CD4^+$ T cells and irradiated PBMC from healthy donors were incubated at the upper chamber on a cell-impermeable polyester membrane with 0.4 μM

pores. The irradiated PBMCs are necessary for optimal PHA activation but die during the culture period, leaving the upper chamber containing a pure population of patient $CD4^+$ T cells. CD8-depleted allogeneic lymphoblasts were added to the lower chamber of the transwell culture in the same numbers as for standard cultures.

After 21 days of culture, 180 μL of culture supernatant were subjected to p24 ELISA (Perkin Elmer). P24 ELISA >6.25 ng/μL were determined as positive.

Results

To analyze proviruses that did not give rise to infectious virus in the viral outgrowth assay (non-induced proviruses), we first established that the conditions were sufficient to activate 100% of resting $CD4^+$ T cells. Resting $CD4^+$ T cells from patients on suppressive ART for >6 months (range, 19–108) were labeled with carboxyfluorescein succinimidyl ester (CFSE) and stimulated with PHA and irradiated allogeneic PBMC under conditions used in the viral outgrowth assay. By day 7, >99.8% of patient cells had divided at least once (Figure 1A), confirming that PHA causes uniform T cell activation.

In the viral outgrowth assay, viruses released after reversal of latency replicate in healthy donor $CD4^+$ lymphoblasts added to the cultures. To facilitate cloning of non-induced proviruses, we explored whether comparable levels of activation and viral outgrowth could be achieved in transwell cultures in which patient cells were separated from donor lymphoblasts by a cell-impermeable membrane (Figure 1B). In side-by-side comparison with standard viral outgrowth assay cultures from ten patients, transwell cultures showed comparable cellular activation in both p24 positive and negative wells, as >95% of patient cells expressed HLA-DR and/or CD25 on day 21 (Figure 1C).

Transwell cultures also showed viral outgrowth comparable to that observed for standard viral outgrowth assay cultures (Figure 1D). Non-induced proviruses were thus cloned from p24-negative wells of limiting dilution transwell and standard viral outgrowth assay cultures.

Discussion

Although difficult and time-consuming, the viral outgrowth assay (Eriksson et al., 2013; Finzi et al., 1997), which has recently been simplified (Laird et al., 2013), does allow detection of cells harboring replication-competent virus. We first explored to make sure this is not simply an issue of assay sensitivity. We showed that the PHA stimulation activates all resting $CD4^+$ T cells as assessed by proliferation and activation marker expression. Using prolonged culture and sensitive RT-PCR assays, we also verified that wells from which non-induced proviruses were obtained were truly negative for viral outgrowth. It is also unlikely that these p24 negative wells remained negative because of reduced viral fitness, as we showed that intact non-induced proviruses had growth kinetics comparable to induced viruses from p24 positive wells (see below). Taken together, these results confirm that we are examining a population of intact proviruses that were not induced to produce infectious virus after a single round of maximum *in vitro* activation.

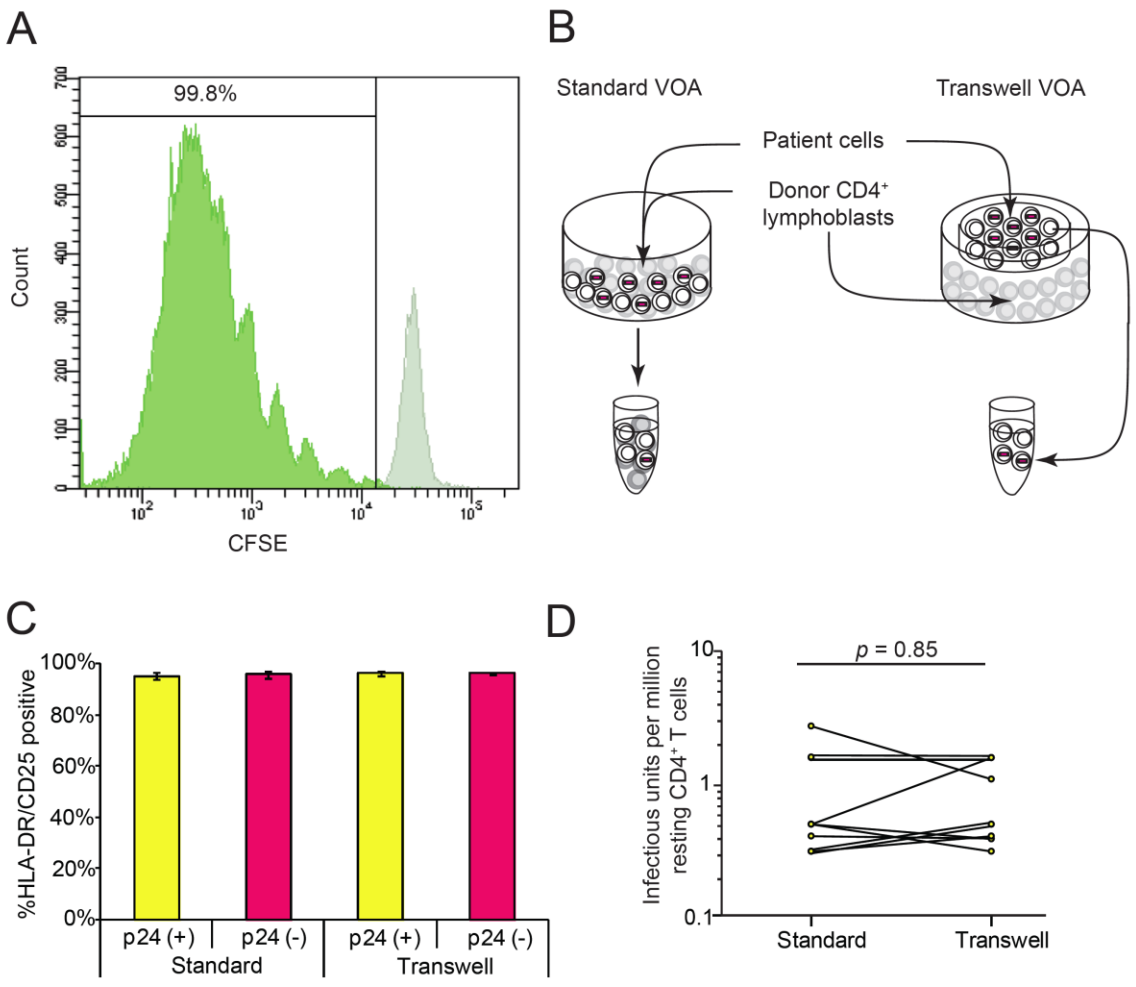
Figure 1. Viral outgrowth assay conditions achieve maximum *in vitro* activation and outgrowth

(A) CFSE fluorescence in labeled patient resting $CD4^+$ T cells on day 0 (light green) and day seven (dark green) of the viral outgrowth assay.

(B) Separation of patient cells from healthy donor $CD4^+$ lymphoblasts in a transwell version of the viral outgrowth assay. Patient cells (white) were activated with PHA and irradiated allogeneic PBMC (not shown). To expand the viruses released from infected cells, $CD4^+$ T lymphoblasts from healthy donors (grey) were added directly to the culture (standard viral outgrowth assay) or were separated from patient cells by a cell-impermeable polyester membrane with 0.4 μm pores (transwell viral outgrowth assay).

(C) Comparison of activation status of patient cells in standard and transwell viral outgrowth assay cultures, and in p24 ELISA positive and negative wells on day 21. HLA-A2 negative patient cells were co-cultured with HLA-A2 positive healthy donor $CD4^+$ lymphoblasts and irradiated allogeneic PBMC. The viable HLA-A2 negative population was gated as patient cells for surface activation marker analysis.

(D) Comparison of viral outgrowth assay results by p24 ELISA of supernatants from parallel standard and transwell viral outgrowth assay cultures from ten patients.


A
99.8%
Count
CFSE
B
Standard VOA
Transwell VOA
Patient cells
Donor CD4+ lymphoblasts
C
%HLA-DR/CD25 positive
100%
80%
60%
40%
20%
0%
p24 (+)
p24 (-)
p24 (+)
p24 (-)
Standard
Transwell
D
p = 0.85
Infectious units per million resting CD4+ T cells
10
1
0.1
Standard
Transwell

Chapter 3. Characterization of non-induced proviral clones

Introduction

The non-induced proviruses had been considered to be defective through various mechanisms, including apolipoprotein B mRNA editing enzyme (APOBEC) mediated G-to-A hypermutations, recombination during reverse transcription (Sanchez et al., 1997; Temin, 1993) and low fidelity of reverse transcriptase (3.4×10^{-5} mutations/bp/cycle (Mansky and Temin, 1995)). The incoming virus is infectious and intact upon entry, but becomes defective during reverse transcription. As long as the very ends of the reverse-transcribed proviral DNA are intact, these defective proviruses will be integrated into the human genome. Based on the analysis of proviruses from acutely infected patients, at least one third are overtly defective (Salazar-Gonzalez et al., 2009), not including large internal deletions. Cells harboring these defective proviruses, which will not produce viral proteins to be presented and recognized by CTLs, continue to expand through homeostatic proliferation. However, these defective proviral genomes remain detectable by PCR-based methods, and are being used in evaluating eradication trials (Katlama et al., 2013). Characterization of these non-induced proviruses is essential in the eradication effort, to avoid prolonged patient exposure to toxic latency reversing agents for overestimating the latent reservoir by detection of defective "graveyard" proviruses.

The main barrier to identify full-length proviral clones is the rarity of the proviruses in resting $CD4^+$ T cells. Identification of full-length HIV-1 genome has been reported, but all from viremic patients with predominant clones in the circulation (Ehrenberg and Michael, 2005; Li et al., 2007; Li et al., 1991; Sahu et al., 2010; Salazar-

Gonzalez et al., 2009). Identification of clonal non-induced proviral clones requires limiting dilution PCR to avoid complementing different proviral clones in one reaction, causing a hybrid of different proviral genomes. Full-length proviral PCR amplification of single clones has been inefficient and challenging, due to the sequence variability in different patients and the bulk of the human genome in PCR reactions. We here established a near-full-length nested PCR to identify single clones of non-induced proviruses and characterized the defects in the proviral genome.

Methods

Study subjects

Peripheral blood was obtained from healthy volunteers and HIV-1 infected donors (Table 1) who had suppression of viremia to <50 copies HIV-1 RNA/ml for >6 months on ART to avoid preintegration latency(Blankson et al., 2000). This study was approved by the Johns Hopkins Institutional Review Board. Written informed consent was obtained from all participants.

Characterization of full-length non-induced proviruses

Genomic DNA was extracted from cells of p24 negative wells using Gentra Puregene Cell Kit (Qiagen). Genomic DNA isolated from p24 negative wells seeded with 4×10^4 or 2×10^5 patient resting $CD4^+$ T cells was subjected to limiting dilution prior to amplification with an initial near full genome length outer PCR (U5 to U5) followed nested amplification of a segment of the *gag* gene. Aliquots were serially diluted and distributed into 96 well plates. Each well was filled up to a final volume of 50 μL with

PCR reaction mixture solutions including outer PCR primers (BLOuterF and BLOuterR). Touchdown cycling conditions with decreasing annealing temperature were used to increase specificity (Table 2). Then, 2 μL aliquots from each outer PCR well were subjected to nested *gag* PCR and 1% agarose gel electrophoresis. If <20% of the *gag* PCR wells were positive, the corresponding outer PCR wells contained one template with >90% of possibility, based on the Poisson distribution. From these corresponding positive outer PCR wells of such plates, 2 μL aliquots were taken for each of four nested PCRs (A, B, C and D). The products were directly sequenced without cloning. Sequences were aligned using CodonCode Sequence Assembly and Alignment Software. The presence of double peaks on sequencing chromatograms was taken as evidence that more than one template may have been present initially, and the sequence was discarded. Direct sequencing results were analyzed by the Los Alamos Hypermut program and HIVAlign program to identify hypermutation and nonsense mutations.

Results

Clonal amplification and sequencing of non-induced proviruses

We obtained near full-length clonal sequences of non-induced proviruses from eight patients on suppressive ART. Patient characteristics are in Table 1. Non-induced proviruses were obtained from wells that were seeded with 4×10^4 or 2×10^5 resting $CD4^+$ T cells and that were negative for p24 on day 21. In clonal viral outgrowth assay cultures, wells with replicating virus are p24-positive by day 10 – 14 (Laird et al., 2013). Even with a more sensitive RT-PCR assay for HIV-1 mRNA (Laird et al., 2013), none of

the p24 negative wells showed exponential growth. Thus, the non-induced proviruses were obtained from wells with no replicating virus despite maximal T cell activation.

Non-induced proviruses were amplified in limiting dilution PCRs to avoid *in vitro* recombination. A near full-length 9.1 kb outer PCR (Li et al., 2007) spanning U5 to U5 (positions 623 – 9,686 by HXB2 coordinates) was followed by nested inner PCRs. Aliquots from the outer PCR were first amplified in a nested inner *gag* PCR (Figure 2A). Cell dilutions for which <20% of the inner reactions were positive were selected because positive wells at these dilutions have >90% probability of being clonal. Aliquots from the *gag*-positive, clonal outer PCRs were amplified using four sets of inner PCR primers to obtain fragments overlapping by 150 – 3,173 bp (Figure 2A). Importantly, instead of cloning PCR products, we directly sequenced them. This dramatically reduces PCR errors since errors occurring after the first or second cycle are present in too small a fraction of the final products to be observed. Sequences with double peaks or non-identical overlap regions were discarded. We identified 213 non-induced proviruses from eight patients.

Hypermutation and large internal deletions render most non-induced proviruses defective

Most (88.3%) non-induced proviruses had obvious defects precluding replication (Figure 1C). Direct sequencing of the nested *gag* PCR product revealed that ~1/3 (32.4%) of non-induced proviruses had APOBEC3G-mediated G→A hypermutation occurring in the expected sequence context (GG or GGG) (Yu et al., 2004), with the first G being mutated. Hypermutated proviruses are replication-defective due to start codon mutations

and numerous tryptophan→stop codon mutations (Figure 3). Although the *gag* gene was analyzed here, other regions of the genome show even greater hypermutation (Yu et al., 2004). Importantly, it is unlikely that hypermutated proviruses could produce functional viral proteins due to stop codons in most open reading frames (ORFs).

Non-induced proviruses that were not hypermutated were further analyzed by nested amplification of four overlapping subgenomic fragments (Figure 2A). Of the 144 clonal non-induced proviruses that were not hypermutated, 97 had large internal deletions identified by smaller amplicon size in agarose gel electrophoresis (Figure 2B). We mapped the precise deletion junctions in 58 of these clones (Figure 4A). For example, clone 10CB7_48H1 (Figure 2B) gave a smaller amplicon for the nested C reaction, and amplification of fragments A, B, and D failed due to deletion of nucleotides 4,869 – 9,533. All 58 mapped deletions would affect expression of the essential regulatory proteins Tat and Rev (Figure 4A) because the deletions encompass the *tat* and *rev* exons, the splice sites, and the Rev-responsive element (RRE).

Deletions are not unexpected. HIV-1 is prone to recombination due to pseudodiploidity (two RNA copies per virion with physical proximity for recombination). Frequent template switching events occur during reverse transcription (Simon-Loriere and Holmes, 2011). Switching between short repeats in a single genome results in deletion of the intervening sequence and one repeat (Temin, 1993). Frequent large deletions have been observed in unfractionated PBMC from viremic patients (Sanchez et al., 1997). Several lines of evidence in our analysis suggest that these deletions occur *in vivo* rather than during *in vitro* analysis. First, deletions were observed following direct sequencing of uncloned PCR products. Second, for a given provirus, precisely the same

deletion junctions were observed in different nested PCRs using different primers. Third, short amplicons were not seen in control experiments with plasmids carrying the reference proviral genomes NL4-3 and BaL. Plasmids were mixed, diluted to eight copies per 100,000 human genome equivalents, and amplified under the same conditions. No deletion or recombination was observed. Fourth, short sequence repeats were identified at some deletion junctions (Figure 2B), consistent with a single polymerase jump due to copy choice recombination during reverse transcription of the minus strand (Sanchez et al., 1997). Taken together, these results demonstrate that a large fraction of non-induced proviruses are non-functional due to large internal deletions, likely introduced during reverse transcription.

The precise fraction of proviruses with deletions could be underestimated by this analysis because deletions could affect PCR primer binding sites. For 39 clones, the exact deletion junction could not be identified, probably because the deletions encompassed some of the binding sites for primers used in the nested PCRs. For these clones, we obtained at least two sequences to ensure that the amplicons contained non-hypermutated, patient-specific sequences. For 12 mapped deletions, the deletion included the reverse *gag* primer binding site, but mapping was possible because other nested reactions were successful.

Mutations in cis elements render some non-induced proviruses defective

Proviruses with correct amplicon size were directly sequenced. A small fraction (8/213) had non-sense mutations and/or frame-shifting insertions or deletions in one or more ORFs. Small deletions (8 – 98 bp) were found in the packaging signal (Ψ) in 12

sequences (Figure 5A). The deletions encompassed the major splice donor (MSD) site. Point mutations in the MSD site were found in two proviruses. Full genome sequencing showed that seven of the non-induced proviruses with Ψ or MSD mutations were otherwise intact. To determine whether these mutations rendered non-induced proviruses defective, we reconstructed three clones by genome synthesis as described below. The reconstructed proviruses included two clones with short (8 bp and 16 bp) Ψ deletions in packaging stem loop two and one with a MSD site mutation (TG|GT→TG|GG). Although these clones had intact ORFs, they did not replicate in healthy donor $CD4^+$ lymphoblasts (Figure 5B) under conditions in which other reconstructed proviruses replicated well (see below). Thus mutations in *cis* elements render otherwise intact proviruses defective.

Discussion

We show here that most non-induced proviruses were rendered defective during reverse transcription by APOBEC3G-induced hypermutation (Yu et al., 2004), by internal deletions caused by copy choice recombination during reverse transcription (Sanchez et al., 1997), or by frame-shift or nonsense mutations caused by the error-prone reverse transcriptase (Bebenek et al., 1989). The resulting defective viral genomes can still integrate because only defects at the ends of the genome affect integration. The defective genomes will be detected in most PCR-based assays of proviral DNA provided that the primer binding sites are intact. Many of the defective proviruses have large internal deletions encompassing the Tat and Rev ORFs and the RRE. Tat-mediated transactivation is required for effective transcriptional elongation (Kao et al., 1987) and

the production of virus particles requires that singly spliced and unspliced HIV-1 mRNAs be exported from the nucleus in a Rev-dependent fashion (Malim et al., 1989). Thus these deleted proviruses may not produce viral proteins even after successful induction of transcription. The same is true for hypermutated proviruses, which have stop codons in every ORF. Importantly, eradication strategies depend on the production of viral proteins which allows recognition of the infected cells by HIV-1 specific CTL (Shan et al., 2012). Defective proviruses with large internal deletions and/or APOBEC3G-induced hypermutation may not be eliminated even by strategies that effectively eliminate cells carrying replication-competent virus. These considerations highlight the difficulty of assessing eradication strategies with PCR-based assays.

Table 1. Patient Characteristics

Patient ID #	Age	Gender	Ethnicity[a]	Time before viral load <50[b] (months)	Time with viral load <50[c] (months)	Regimen on enrollment[d]
9	57	M	AA	19	28	TDF/FTC, FPV/r
10	47	M	W	66	60	TDF/FTC, ATZ/r
16	59	M	AA	51	33	FTC, EFV, ATZ/r
17	49	M	AA	141	57	ZDV/ 3TC, EFV, ATZ/r
19	46	F	AA	45	105	ABC/3TC, NVP
20	75	F	AA	81	88	ABC/3TC/ZDV, TDF, LPV/r
22	38	M	AA	6	52	ABC/3TC/EFV
23	38	M	AA	118	108	ABC/3TC/EFV

[a]AA, African-American; W, White.
[b]Time after diagnosis of infection before documented period of sustained suppression of plasma HIV-1 RNA to <50 copies/mL on ART
[c]Time of continuous suppression of plasma HIV-1 RNA to <50 copies/mL prior to enrollment
[d]Drug abbreviations: ABC, abacavir; ATZ, atazanavir; EFV, efavirenz; FPV, fosaprenavir; FTC, emtricitabine; LPV, lopinavir; NVP, nevirapine; TDF, tenofovir disoproxil fumarate; ZDV, zidovudine; 3TC, lamivudine; /r, boosted with ritonavir.

Table 2. PCR primers and conditions

PCR	Length	Name	HXB2 position	Exact sequence	Extension time
Outer PCR					
	9,064	BLOuterF	623 – 649	AAATCTCTAGCAGTGGCGCCCGAACAG	10 m
		BLOuterR	9,662 – 9,686	TGAGGGATCTCTAGTTACCAGAGTC	
Inner PCR					
A	4,449	275F	646 – 666	ACAGGGACCTGAAAGCGAAAG	5 m
		3INOut	5,072 – 5,094	AATCCTCATCCTGTCTACTTGCC	
B	5,793	263F	651 – 672	GACCTGAAAGCGAAAGGGAAAC	6 m
		3AccOut	6,421 – 6,443	GGCATGTGTGGCCCARAYATTAT	
C	6,385	5INOut	3,248 – 3,270	ACTCCATCCTGATAAATGGACAG	6 m 30 s
		BLInnerR	9,604 – 9,632	GCACTCAAGGCAAGCTTTATTGAGGCTTA	
D	4,778	5AccOut	4,899 – 4,922	CGGGTTTATTACAGGGACARCARA	5 m
		280R	9,650 – 9,676	CTAGTTACCAGAGTCACACAACAGACG	
gag PCR for clonality					
gag	1,448	5GagIn	836 – 857	GGGAAAAAATTCGGTTAAGGCC	1 m 30 s
		3GagIn	2,264 – 2,283	CGAGGGGTCGTTGCCAAAGA	

Outer PCR
Platinum *Taq* High Fidelity Polymerase (Invitrogen) was used for all reactions.
94°C for 2 m; then 94°C for 30 s, 64°C for 30 s, 68°C for 10 m for 3 cycles; 94°C for 30 s, 61°C for 30 s, 68°C for 10 m for 3 cycle; 94°C for 30 s, 58°C for 30 s, 68°C for 10 m for 3 cycle; 94°C for 30 s, 55°C for 30 s, 68°C for 10 m for 41 cycle; then 68°C for 10 m.

Inner and *gag* PCR cycling condition were the same as for the outer PCR but with different extension times as indicated.

Figure 2. Characterization of non-induced proviruses

(A) Strategy and results of analysis of non-induced proviruses. Limiting dilution viral outgrowth assay cultures were established from eight patients. A total of 51 p24-negative wells containing a total of 8.9 x 10^6 resting $CD4^+$ T cells were analyzed. 213 non-induced proviruses were identified by near-full length limiting dilution PCR followed by a nested *gag* PCR. The *gag* PCR products were directly sequenced, and proviruses with APOBEC3G-mediated G to A hypermutation were identified using the Los Alamos Hypermut algorithm (Rose and Korber, 2000). Non-hypermutated proviruses were analyzed by nested amplification of four overlapping subgenomic fragments (labeled A–D). Amplicons with patient HIV-1 sequence but discernibly smaller than expected size in 1% agarose gel electrophoresis were considered to contain large internal deletions which, where possible, were mapped by direct sequencing. Other lethal defects including single nucleotide insertions and deletions (INDELs), packaging signal (Ψ) deletions, and major splice donor (MSD) site mutations were identified by direct sequencing.

(B) Results of 1% agarose gel electrophoresis of four subgenomic amplicons (A – D) from representative non-induced proviruses with different defects. Left to right: intact genome, 16 bp Ψ deletion, deletion with defined junction, and deletion with undefined junction. Clone name is give below each gel. Triangles, amplicons with correct size. Asterisks, amplicons with smaller than expected size.

(C) Summary of the analysis of 213 non-induced proviruses.

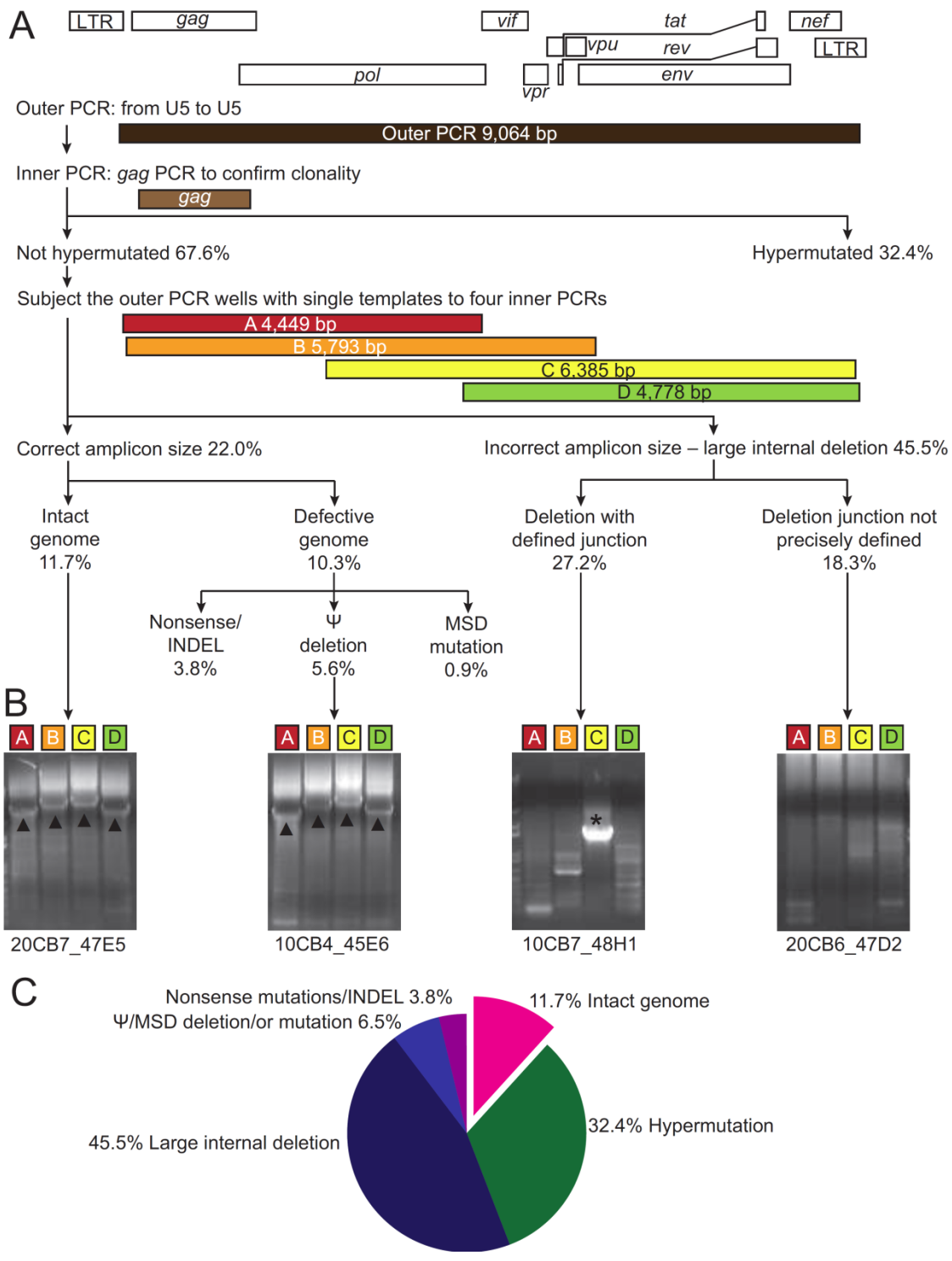
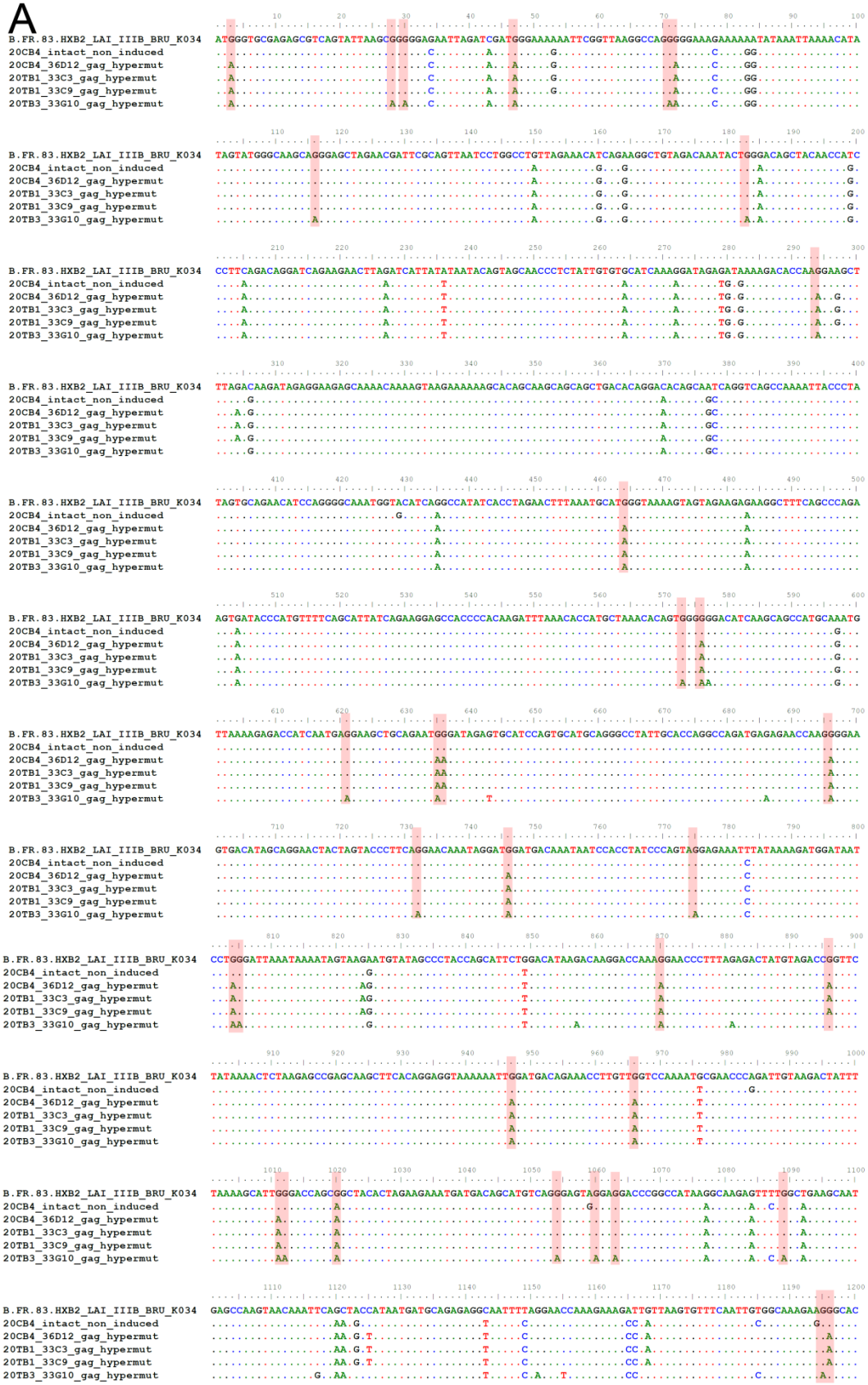

A
LTR
gag
vif
tat
nef
vpu
rev
LTR
pol
vpr
env
Outer PCR: from U5 to U5
Outer PCR 9,064 bp
Inner PCR: gag PCR to confirm clonality
gag
Not hypermutated 67.6%
Hypermutated 32.4%
Subject the outer PCR wells with single templates to four inner PCRs
A 4,449 bp
B 5,793 bp
C 6,385 bp
D 4,778 bp
Correct amplicon size 22.0%
Incorrect amplicon size – large internal deletion 45.5%
Intact genome 11.7%
Defective genome 10.3%
Deletion with defined junction 27.2%
Deletion junction not precisely defined 18.3%
Nonsense/ INDEL 3.8%
Ψ deletion 5.6%
MSD mutation 0.9%
B
A B C D
20CB7_47E5
10CB4_45E6
10CB7_48H1
20CB6_47D2
C
Nonsense mutations/INDEL 3.8%
Ψ/MSD deletion/or mutation 6.5%
11.7% Intact genome
32.4% Hypermutation
45.5% Large internal deletion

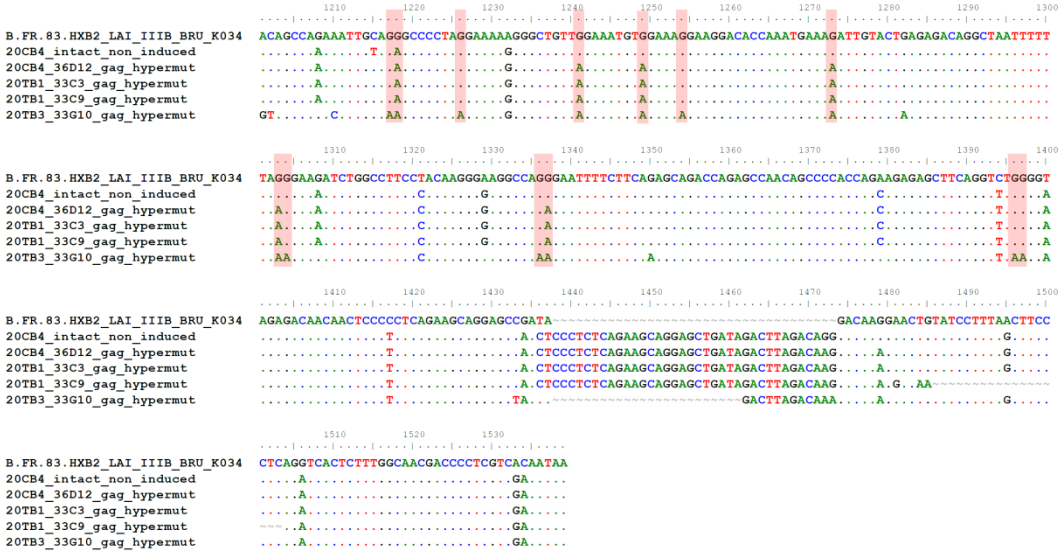
Figure 3. **APOBEC3G-mediated G to A hypermutation renders some non-induced proviruses defective**

(A) Nucleotide sequences of the *gag* gene (representing nucleotides 790 – 2,292, HXB2 coordinates) of hypermutated non-induced proviruses from a representative patient (#20). Sequence changes representing probable G to A hypermutation are shaded. Sequences are aligned to the reference sequence HXB2 and an intact, non-induced provirus from the same patient.

(B) Amino acid sequences of Gag from the hypermutated non-induced proviruses from (A). Sequence changes resulting from probable G to A hypermutation are shaded. Sequences are aligned to the reference sequence HXB2 and an intact, non-induced provirus from the same patient. Note that the methionine start codon of Gag is frequently mutated to isoleucine (ATG → ATA). This is because the second amino acid of Gag is glycine (GGX), which results in an APOBEC3G recognition site AT<u>GG</u>GX, allowing mutation of the third position of the start codon. Hypermutation also affects many tryptophan codons which contain the APOBEC3G recognition site. Mutation of either or both of the G nucleotides to A generates a stop codon (TGG → TGA, TAG, or TAA, designated with an *).

A


B.FR.83.HXB2_LAI_IIIB_BRU_K034
20CB4_intact_non_induced
20CB4_36D12_gag_hypermut
20TB1_33C3_gag_hypermut
20TB1_33C9_gag_hypermut
20TB3_33G10_gag_hypermut


B.FR.83.HXB2_LAI_IIIB_BRU_K034
20CB4_intact_non_induced
20CB4_36D12_gag_hypermut
20TB1_33C3_gag_hypermut
20TB1_33C9_gag_hypermut
20TB3_33G10_gag_hypermut

B

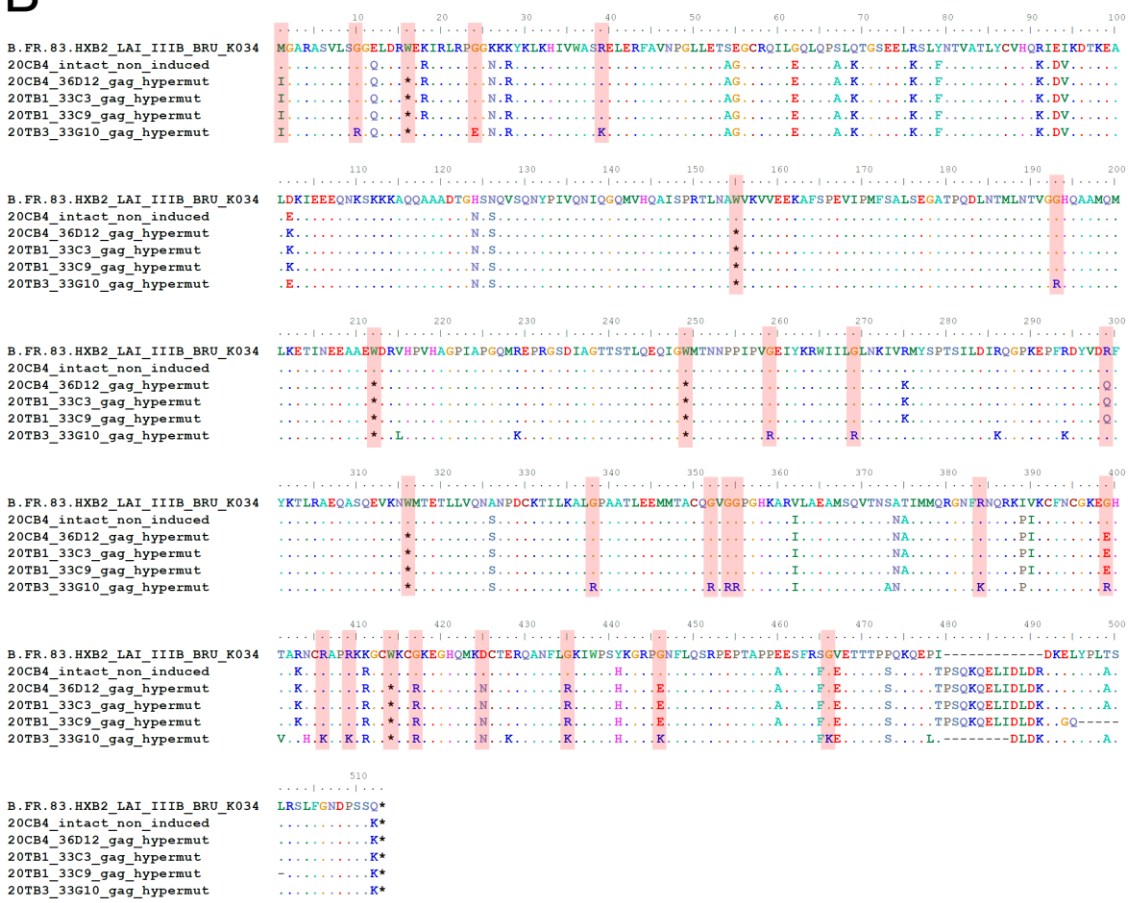
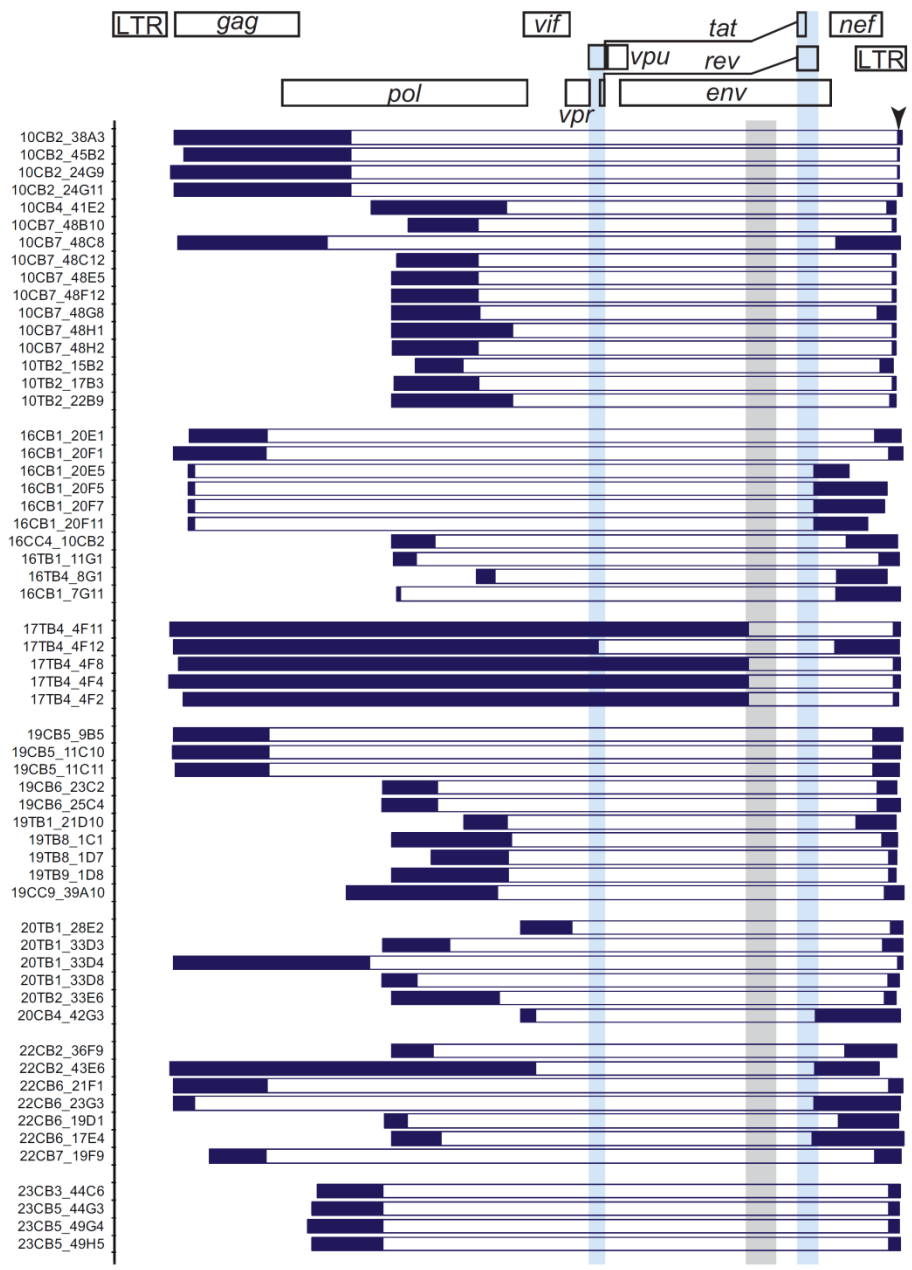

B.FR.83.HXB2_LAI_IIIB_BRU_K034
20CB4_intact_non_induced
20CB4_36D12_gag_hypermut
20TB1_33C3_gag_hypermut
20TB1_33C9_gag_hypermut
20TB3_33G10_gag_hypermut

Figure 4. Mapping of large internal deletions in non-induced proviruses

(A) Locations of deletions. Dark blue horizontal bars: continuous sequencing reads interrupted by deletions. White horizontal bars: deletions. Light blue vertical bars, *tat* and *rev* ORFs. Gray vertical bar: RRE element.

(B) Representative sequences at deletion junctions. Short repeats (underlined) at the deletion junctions suggest a single reverse transcriptase jump due to copy choice recombination (Sanchez et al., 1997; Temin, 1993). Hypermutated sequences (see Figure S2) were not analyzed for deletions.

A


LTR
gag
vif
tat
nef
vpu
rev
LTR
pol
vpr
env
10CB2_38A3
10CB2_45B2
10CB2_24G9
10CB2_24G11
10CB4_41E2
10CB7_48B10
10CB7_48C8
10CB7_48C12
10CB7_48E5
10CB7_48F12
10CB7_48G8
10CB7_48H1
10CB7_48H2
10TB2_15B2
10TB2_17B3
10TB2_22B9
16CB1_20E1
16CB1_20F1
16CB1_20E5
16CB1_20F5
16CB1_20F7
16CB1_20F11
16CC4_10CB2
16TB1_11G1
16TB4_8G1
16CB1_7G11
17TB4_4F11
17TB4_4F12
17TB4_4F8
17TB4_4F4
17TB4_4F2
19CB5_9B5
19CB5_11C10
19CB5_11C11
19CB6_23C2
19CB6_25C4
19TB1_21D10
19TB8_1C1
19TB8_1D7
19TB9_1D8
19CC9_39A10
20TB1_28E2
20TB1_33D3
20TB1_33D4
20TB1_33D8
20TB2_33E6
20CB4_42G3
22CB2_36F9
22CB2_43E6
22CB6_21F1
22CB6_23G3
22CB6_19D1
22CB6_17E4
22CB7_19F9
23CB3_44C6
23CB5_44G3
23CB5_49G4
23CB5_49H5

B

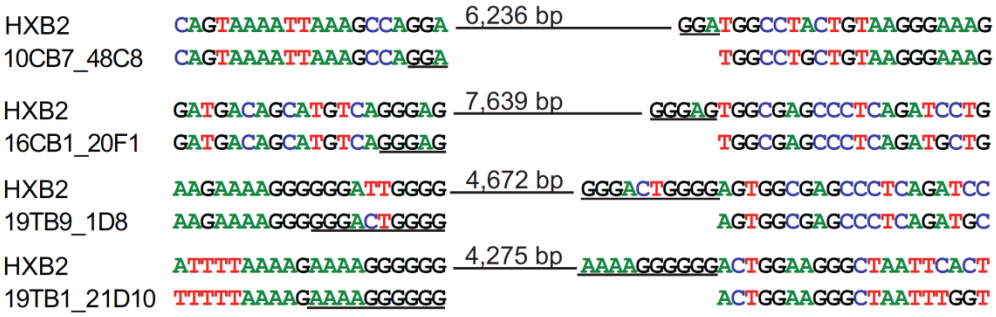
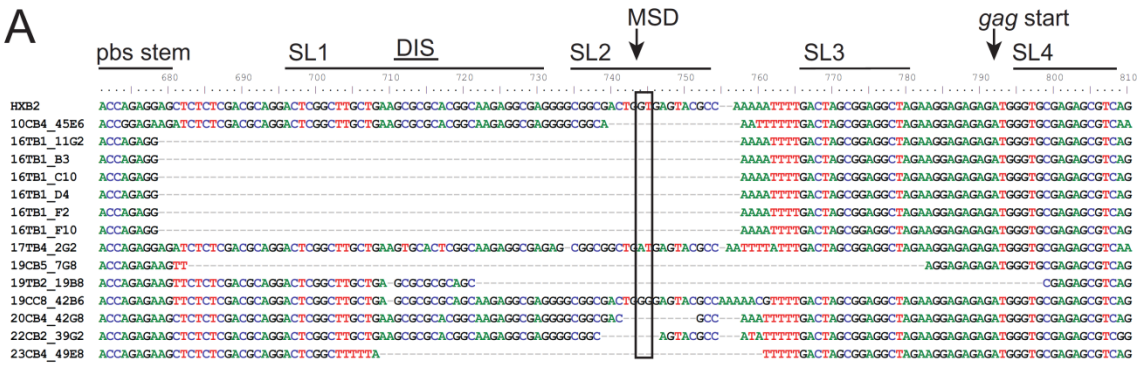

HXB2
CAGTAAAATTAAAGCCAGGA
6,236 bp
GGATGGCCTACTGTAAGGGAAAG
10CB7_48C8
CAGTAAAATTAAAGCCAGGA
TGGCCTGCTGTAAGGGAAAG
HXB2
GATGACAGCATGTCAGGGAG
7,639 bp
GGGAGTGGCGAGCCCTCAGATCCTG
16CB1_20F1
GATGACAGCATGTCAGGGAG
TGGCGAGCCCTCAGATGCTG
HXB2
AAGAAAAGGGGGGATTGGGG
4,672 bp
GGGACTGGGGAGTGGCGAGCCCTCAGATCC
19TB9_1D8
AAGAAAAGGGGGGACTGGGG
AGTGGCGAGCCCTCAGATGC
HXB2
ATTTTAAAAGAAAAGGGGGG
4,275 bp
AAAAGGGGGGACTGGAAGGGCTAATTCACT
19TB1_21D10
TTTTTAAAAGAAAAGGGGGG
ACTGGAAGGGCTAATTTGGT

Figure 5. Deletions and mutations in *cis* elements

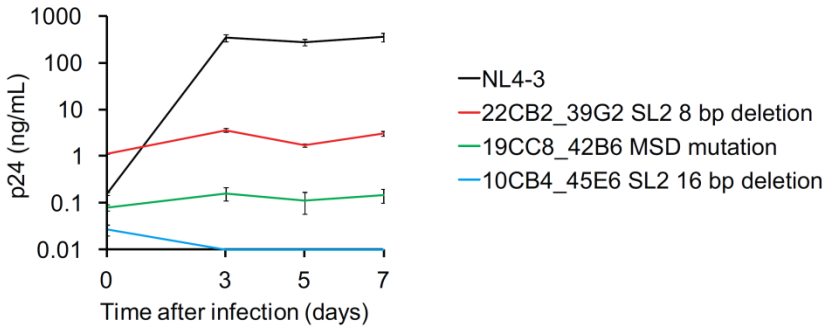
(A) Location of *cis* element deletions and mutations in 14 clones from five patients. pbs, tRNA primer binding site; SL, packaging stem loop; DIS, dimerization initiation site; MSD, major splice donor site, typically TG|**GT** (arrow and box).

(B) Growth kinetics of three reconstructed clones from three patients with MSD site deletion or mutation but intact ORFs for all viral proteins, in comparison with the NL4-3 reference strain. Each viral inoculum was normalized to 200 ng p24/mL.

A


pbs stem
SL1
DIS
SL2
MSD
SL3
gag start
SL4
HXB2
10CB4_45E6
16TB1_11G2
16TB1_B3
16TB1_C10
16TB1_D4
16TB1_F2
16TB1_F10
17TB4_2G2
19CB5_7G8
19TB2_19B8
19CC8_42B6
20CB4_42G8
22CB2_39G2
23CB4_49E8

B


1000
100
10
1
0.1
0.01
p24 (ng/mL)
0
3
5
7
Time after infection (days)
NL4-3
22CB2_39G2 SL2 8 bp deletion
19CC8_42B6 MSD mutation
10CB4_45E6 SL2 16 bp deletion

Chapter 4. Intact non-induced proviruses are replication competent

Introduction

Clonal analysis of full-length HIV-1 genomes from untreated patients has contributed to current understanding of HIV-1 evolution and antiviral immune responses (272 Li,B. 2007; 236 Salazar-Gonzalez,J.F. 2009; 274 Sahu,G.K. 2010; 297 Li,Y. 1991). Reconstruction of near-full-length proviruses had be reported (Gao et al., 1998)(Sahu et al., 2010), but existing reconstruction strategies involve combining two cloned PCR products, which raises the questions of combining different clones into one and introducing PCR errors. Bulk PCR sequencing without cloning will only pick up the true sequences, as PCR induced errors will not be reported in the sequencing results. We therefore choose *de novo* gene synthesis to these avoid cloning induced errors.

The replication competence of the reconstructed proviruses depends not only on intact open reading frames of all viral genomes, but also intact promoter (long terminal repeat, LTR) function and *cis*-acting elements, including the packaging signal (Ψ), *trans*-activation response elements (TAR), ribosomal frameshift signal, polypurine tracks and Rev response element (RRE). The function of these elements, with extensive secondary structures, may not be fully evaluated by sequence analysis alone. Therefore, reconstruction of the full-length genome becomes the only way to evaluate the competence of non-induced proviruses, not only to understand the impact of missense mutations but also the function of the essential elements.

Methods

Reconstruction of full-length non-induced proviruses

Direct sequencing results were used for *de novo* genome synthesis. The resulting sequences were cloned into pNL4-3 vector (GenScript Inc USA and Bio Basic Inc) by *Bss*HII and *Ngo*MIV. To correct the 5'LTR to 100% patient derived, the plasmid fragments between the *Aat*II and *Bss*HII sites were cloned into Zero Blunt TOPO vector (Invitrogen) and subjected to site directed mutagenesis by QuikChange Site-Directed Multi Kit (Agilent Technologies, Inc). Reconstructed plasmids were transformed into Stbl3 chemically competent cells (Invitrogen) and cultured at 30°C to reduce recombination. The reconstructed plasmids were confirmed by restriction digestion and sequencing. Induced proviruses from end-dilution viral outgrowth cultures were reconstructed using similar methods. Plasmids were transfected to HEK 293T cells using Lipofectamine 2000 (Invitrogen). Viruses released into the supernatants were collected by ultracentrifugation 48 – 72 hours after transfection. Viruses were titered and normalized to 200 ng p24/mL for growth kinetics experiments. Healthy donor $CD4^+$ lymphoblasts were infected using spinoculation (1,200 *g,* 25°C for two hours), and growth was monitored by p24 ELISA of culture supernatants.

Measurement of LTR function

Patient-derived LTR sequences were cloned into wt-LTR-Luc reporter and transfected into resting $CD4^+$ T cells from healthy donors using nucleofection (Amaxa) (Yang et al., 2009a). TK-RLuc was used as internal control. After 48 hours, basal luciferase activity was measured. We also measured luciferase activity four hours after stimulation with phorbol myristate acetate (PMA) (10 ng/mL) and ionomycin (1 μM).

Results

Of 213 non-induced proviruses, 25 (11.7%) had intact ORFs and *cis* elements. When sequences were compared with corresponding sequences from induced proviruses from the same patient, no known lethal mutations were seen. To determine whether these *intact non-induced proviruses* are replication-competent, we used the direct sequencing results to reconstruct full-length non-induced proviral clones by *de novo* genome synthesis (Figure 6A). This strategy avoids PCR and cloning-induced errors. We reconstructed six non-induced proviruses from four patients by inserting the synthesized sequence into a plasmid carrying the reference isolate NL4-3. Not captured in our PCR strategy is a 108 bp segment of the provirus, representing nucleotides (nt) 565 – 672 (HXB2 coordinates). This segment in the 5' untranslated region includes part of U5 and the primer binding site (pbs) (Figure 6A). Although U5 deletions may not affect replication-competence (Vicenzi et al., 1994), we took additional steps to make the reconstructed clones fully patient-derived, without any NL4-3 sequence. We used limiting dilution PCR to amplify the LTR-*gag* region from cells in p24 negative wells in separate reactions. Using a 424 bp segment from the 5' U3-R-U5 region (HXB2 nt 140–564), we constructed phylogenetic trees (Figure 7). Then, using site-directed mutagenesis, we corrected the 108 bp segment from NL4-3 to the phylogenetically closest sequence from the same patient (Figure 6A). This process results in proviruses that are 100% patient-derived and 98.2 % equivalent to specific proviruses present *in vivo* with the remaining 1.8% equivalent to a very closely related provirus from the same patient. Given the high sequence conservation in the 108 bp segment (Figure 7), we

estimate that the reconstructed clones could differ from the parent clones by at most 3 nt, or 0.03% of the genome. For each patient, we also reconstructed an induced viral clone from a p24 positive well.

Strikingly, all six reconstructed non-induced proviruses from four different patients showed replication fitness comparable to that of the reference isolate NL4-3 and reconstructed induced proviruses from the same patients (Figure 6B). It is formally possible that these intact non-induced proviruses could have actually been defective, with an inactivating mutation in the 108 bp region that was corrected when the NL4-3 sequence was reverted to the most closely related patient-specific sequence. Such a mutation would have had to arise during reverse transcription in cell from which the proviral sequence was obtained. Based on measured *in vivo* mutation rate of 3.4×10^{-5} mutations/bp/cycle (Mansky and Temin, 1995), the probability of a mutation arising in the 108 bp segment during a single cycle is only 0.0037, and only a subset of these mutations would be inactivating. Importantly, all of the proviruses that we reconstructed were functional. The probability that all of these actually had defects in the 108 bp region is $(0.00324)^6 = 1.16 \times 10^{-15}$. Thus we feel that it is extremely unlikely that the intact non-induced proviruses described here are in fact defective. This conclusion is further supported by experiments showing that an additional round of PHA stimulation causes some non-induced proviruses to produce replication-competent virus (see below). Taken together, these results indicate that a substantial fraction of the non-induced proviruses are intact and capable of generating infectious virus if induced *in vivo*.

Non-induced proviruses have intact promoter function unless hypermutated

The ability of intact non-induced proviruses to produce infectious virus suggests that, at the primary sequence level, their promoters are functional. To confirm this, we cloned LTRs from representative non-induced proviruses into a luciferase reporter construct (Yang et al., 2009a). We measured luciferase activity in transfected resting $CD4^+$ T cells before and four hours after stimulation with PMA and ionomycin. We also analyzed LTRs from induced proviruses from the same patients and NL4-3. In general, LTRs from non-induced proviruses showed basal (Figure 8A, C) and stimulated (Figure 8B, D) activity comparable to LTRs from induced proviruses and NL4-3. Decreased LTR function was only observed for hypermutated clones. This likely reflects G→A hypermutation in binding sites for the transcription factors NF-κB and Sp1 (Figure 8E). Thus, most non-induced proviruses have LTRs that are intact at the primary sequence level.

Figure 6. Growth kinetics of reconstructed non-induced viruses

(A) Reconstruction of full-length non-induced proviruses. Pink, clonal patient-derived sequence. Blue, pNL4-3. Purple, sequence from the most closely related patient-specific virus. See text for details.

(B) Growth kinetics of reconstructed non-induced viruses from p24 negative wells (pink and purple), reconstructed induced viruses from p24 positive wells (yellow), and NL4-3 (black).

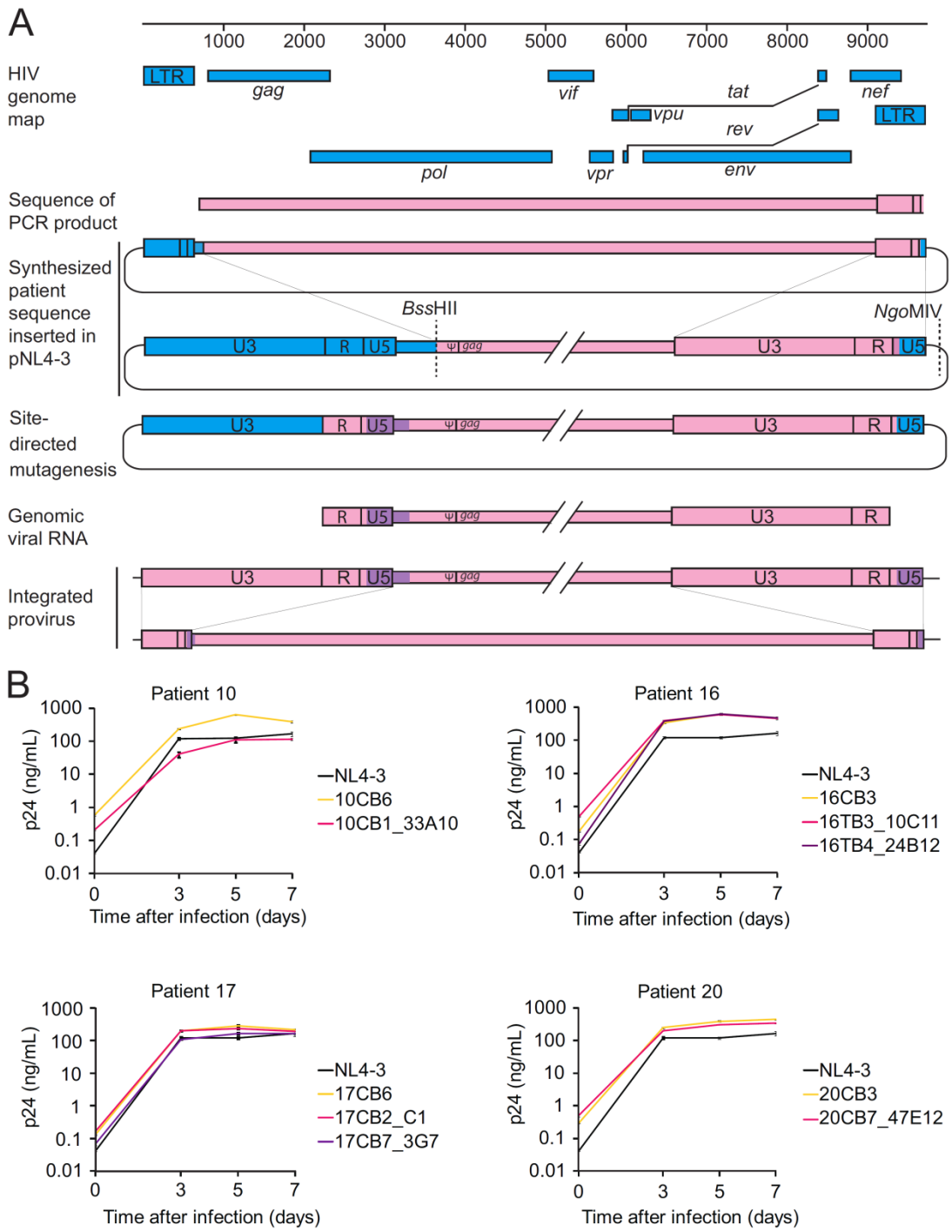
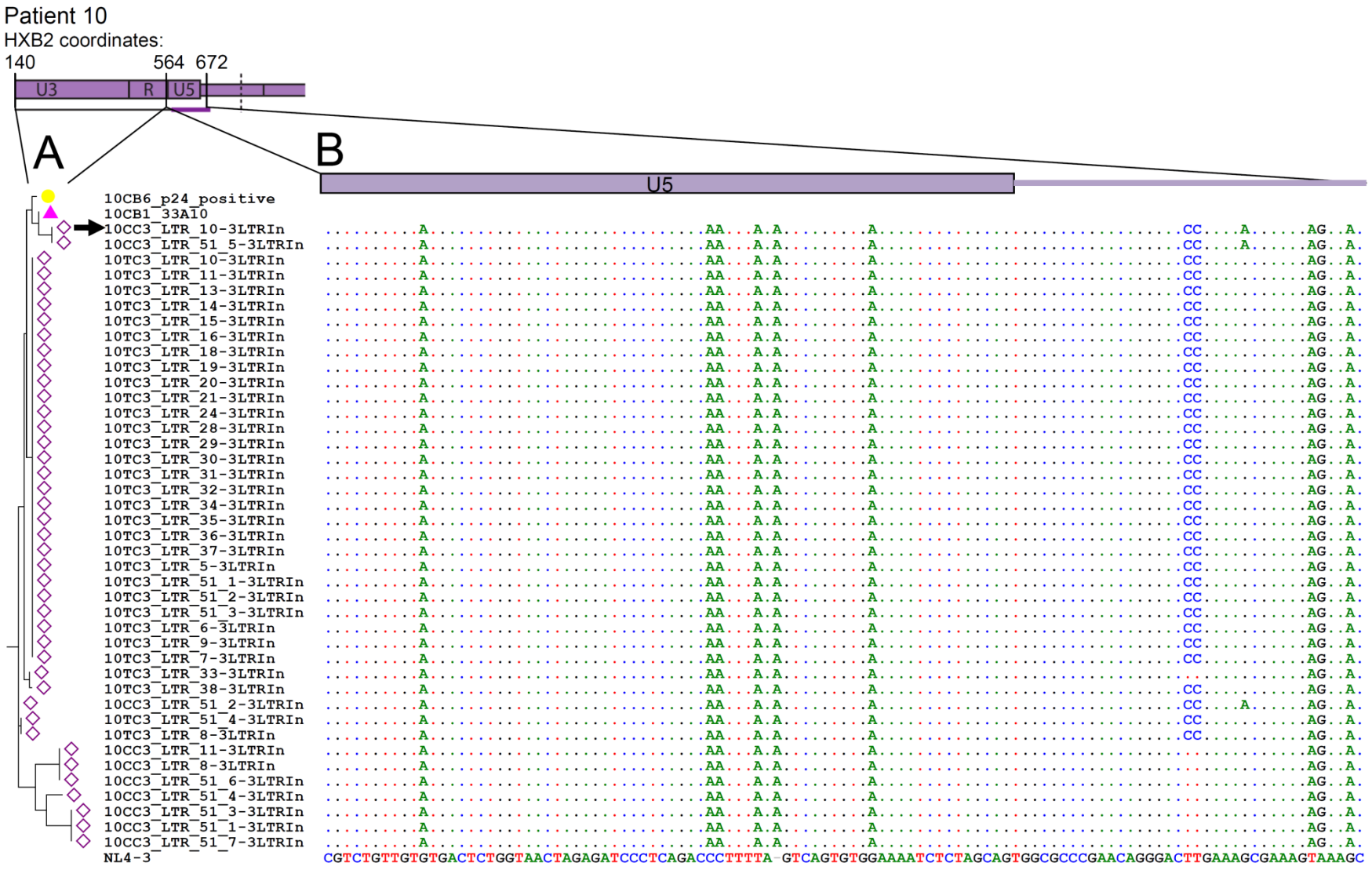
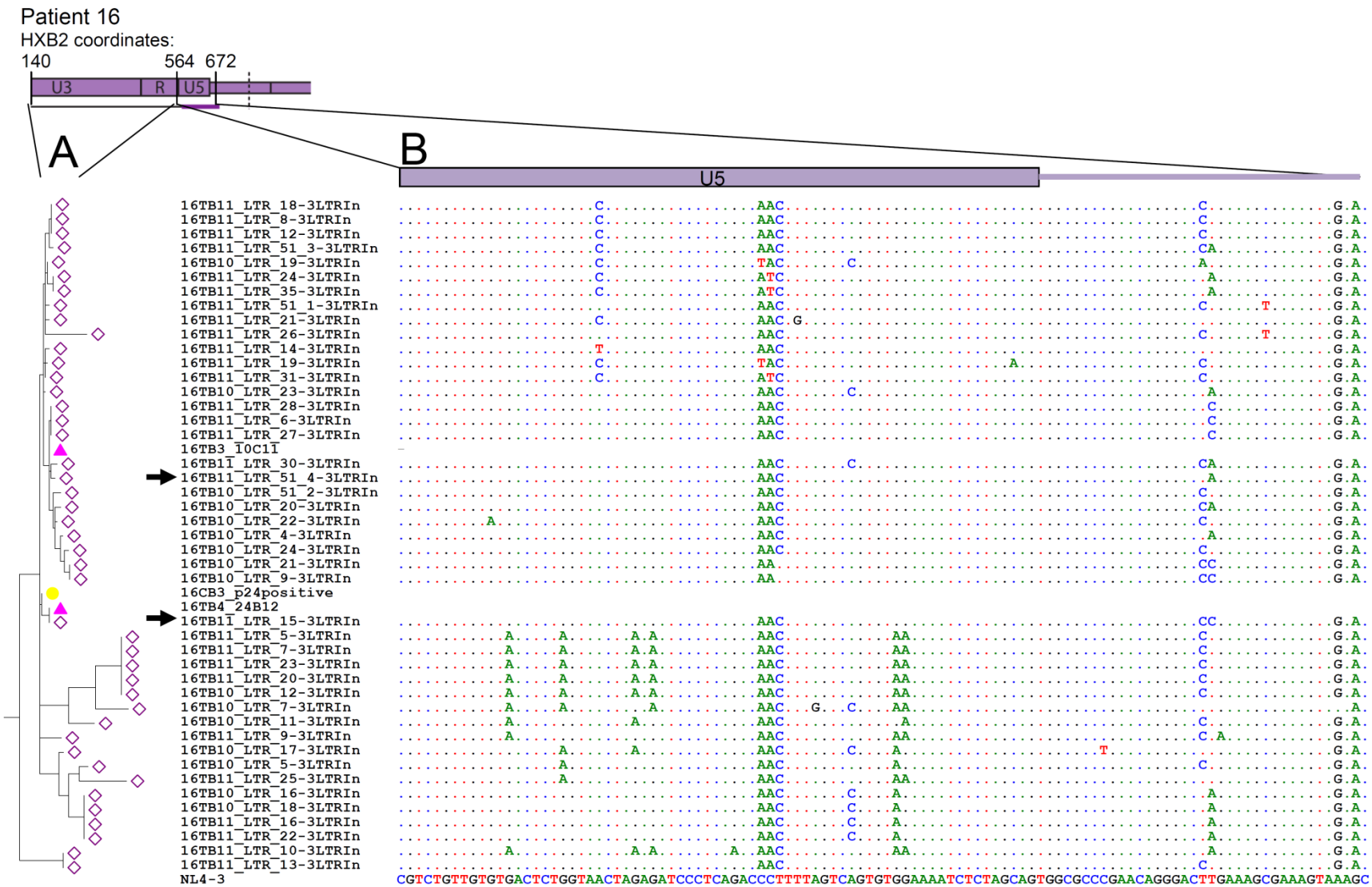

A
1000
2000
3000
4000
5000
6000
7000
8000
9000
HIV genome map
LTR
gag
vif
tat
vpu
rev
nef
LTR
pol
vpr
env
Sequence of PCR product
Synthesized patient sequence inserted in pNL4-3
BssHII
NgoMIV
U3
R
U5
Ψ
gag
Site-directed mutagenesis
Genomic viral RNA
Integrated provirus
B
Patient 10
p24 (ng/mL)
Time after infection (days)
NL4-3
10CB6
10CB1_33A10
Patient 16
16CB3
16TB3_10C11
16TB4_24B12
Patient 17
17CB6
17CB2_C1
17CB7_3G7
Patient 20
20CB3
20CB7_47E12

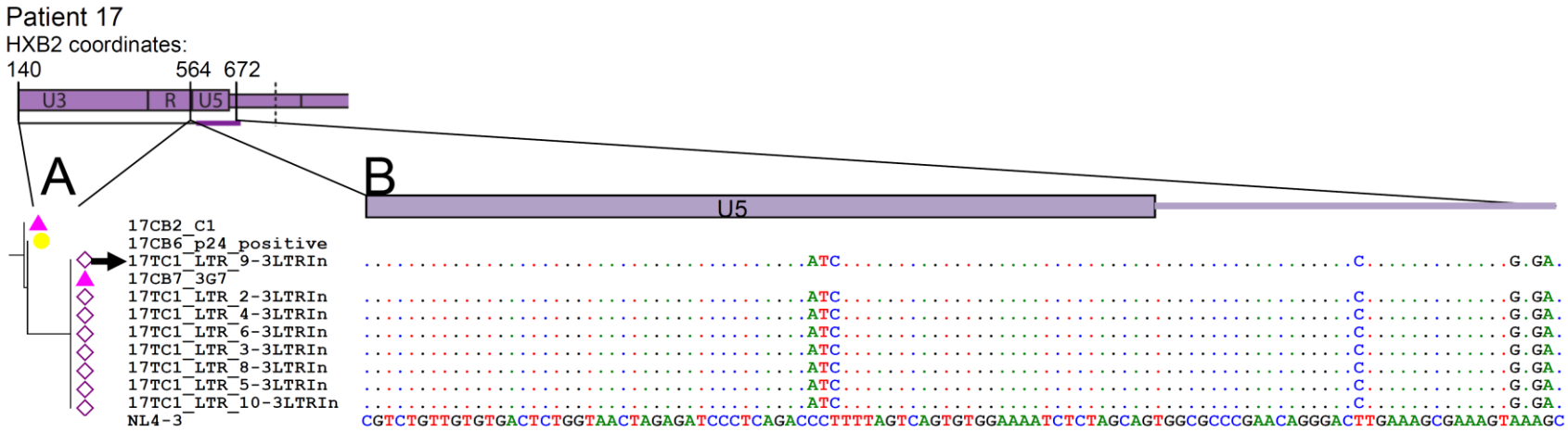
Figure 7. Phylogenetic analysis for reconstruction of the 108 bp sequence

(A) Use of phylogenetic analysis to reconstruct full-length patient-derived proviruses for six non-induced proviral clones and four induced proviral clones. We corrected the 108 bp segment (565 – 672, HXB2 coordinates) missing from the original PCRs (Figure 3A) to the most closely related sequence from the same patient. For this purpose, we used limiting dilution PCR to amplify the LTR-*gag* region from cells of p24 negative wells. An average of 31 clonal sequences were obtained for each patient. Using a 424 bp segment from the U3-R-U5 region (HXB2 positions 140 – 564), we constructed phylogenetic trees and calculated the phylogenetic distance between reconstructed clones and each of these LTR-*gag* PCR sequences using a maximum likelihood estimate. Pink triangles, intact non-induced proviruses to be reconstructed. Yellow circles, induced proviruses from p24 positive wells from the same patient. Open diamonds, non-induced proviral LTR sequences from the same patient. Arrows, exact sequences used for correction, chosen based on smallest pairwise distance in a maximum likelihood phylogenic analysis using MEGA 5.10 software.

(B) Nucleotide sequences of the 108 bp segment (nt 565 – 672). Nucleotide differences between NL4-3 (bottom sequence) and the patient-derived sequence that was most closely related to the reconstructed clone (arrow) were corrected by site directed mutagenesis in the reconstructed clones (blank sequences).


Patient 10
HXB2 coordinates:
140
564
672
U3
R
U5
A
B
U5
10CB6_p24_positive
10CB1_33A10
10CC3_LTR_10-3LTRIn
10CC3_LTR_51_5-3LTRIn
10TC3_LTR_10-3LTRIn
10TC3_LTR_11-3LTRIn
10TC3_LTR_13-3LTRIn
10TC3_LTR_14-3LTRIn
10TC3_LTR_15-3LTRIn
10TC3_LTR_16-3LTRIn
10TC3_LTR_18-3LTRIn
10TC3_LTR_19-3LTRIn
10TC3_LTR_20-3LTRIn
10TC3_LTR_21-3LTRIn
10TC3_LTR_24-3LTRIn
10TC3_LTR_28-3LTRIn
10TC3_LTR_29-3LTRIn
10TC3_LTR_30-3LTRIn
10TC3_LTR_31-3LTRIn
10TC3_LTR_32-3LTRIn
10TC3_LTR_34-3LTRIn
10TC3_LTR_35-3LTRIn
10TC3_LTR_36-3LTRIn
10TC3_LTR_37-3LTRIn
10TC3_LTR_5-3LTRIn
10TC3_LTR_51_1-3LTRIn
10TC3_LTR_51_2-3LTRIn
10TC3_LTR_51_3-3LTRIn
10TC3_LTR_6-3LTRIn
10TC3_LTR_9-3LTRIn
10TC3_LTR_7-3LTRIn
10TC3_LTR_33-3LTRIn
10TC3_LTR_38-3LTRIn
10CC3_LTR_51_2-3LTRIn
10TC3_LTR_51_4-3LTRIn
10TC3_LTR_8-3LTRIn
10CC3_LTR_11-3LTRIn
10CC3_LTR_8-3LTRIn
10CC3_LTR_51_6-3LTRIn
10CC3_LTR_51_4-3LTRIn
10CC3_LTR_51_3-3LTRIn
10CC3_LTR_51_1-3LTRIn
10CC3_LTR_51_7-3LTRIn
NL4-3
CGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTA GTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGC


Patient 16
HXB2 coordinates:
140
564
672
U3
R
U5
A
B
U5
16TB11_LTR_18-3LTRIn
16TB11_LTR_8-3LTRIn
16TB11_LTR_12-3LTRIn
16TB11_LTR_51_3-3LTRIn
16TB10_LTR_19-3LTRIn
16TB11_LTR_24-3LTRIn
16TB11_LTR_35-3LTRIn
16TB11_LTR_51_1-3LTRIn
16TB11_LTR_21-3LTRIn
16TB11_LTR_26-3LTRIn
16TB11_LTR_14-3LTRIn
16TB11_LTR_19-3LTRIn
16TB11_LTR_31-3LTRIn
16TB10_LTR_23-3LTRIn
16TB11_LTR_28-3LTRIn
16TB11_LTR_6-3LTRIn
16TB11_LTR_27-3LTRIn
16TB3_10C11
16TB11_LTR_30-3LTRIn
16TB11_LTR_51_4-3LTRIn
16TB10_LTR_51_2-3LTRIn
16TB10_LTR_20-3LTRIn
16TB10_LTR_22-3LTRIn
16TB10_LTR_4-3LTRIn
16TB10_LTR_24-3LTRIn
16TB10_LTR_21-3LTRIn
16TB10_LTR_9-3LTRIn
16CB3_p24positive
16TB4_24B12
16TB11_LTR_15-3LTRIn
16TB11_LTR_5-3LTRIn
16TB11_LTR_7-3LTRIn
16TB11_LTR_23-3LTRIn
16TB11_LTR_20-3LTRIn
16TB10_LTR_12-3LTRIn
16TB10_LTR_7-3LTRIn
16TB10_LTR_11-3LTRIn
16TB11_LTR_9-3LTRIn
16TB10_LTR_17-3LTRIn
16TB10_LTR_5-3LTRIn
16TB11_LTR_25-3LTRIn
16TB10_LTR_16-3LTRIn
16TB10_LTR_18-3LTRIn
16TB11_LTR_16-3LTRIn
16TB11_LTR_22-3LTRIn
16TB11_LTR_10-3LTRIn
16TB11_LTR_13-3LTRIn
NL4-3
CGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGC

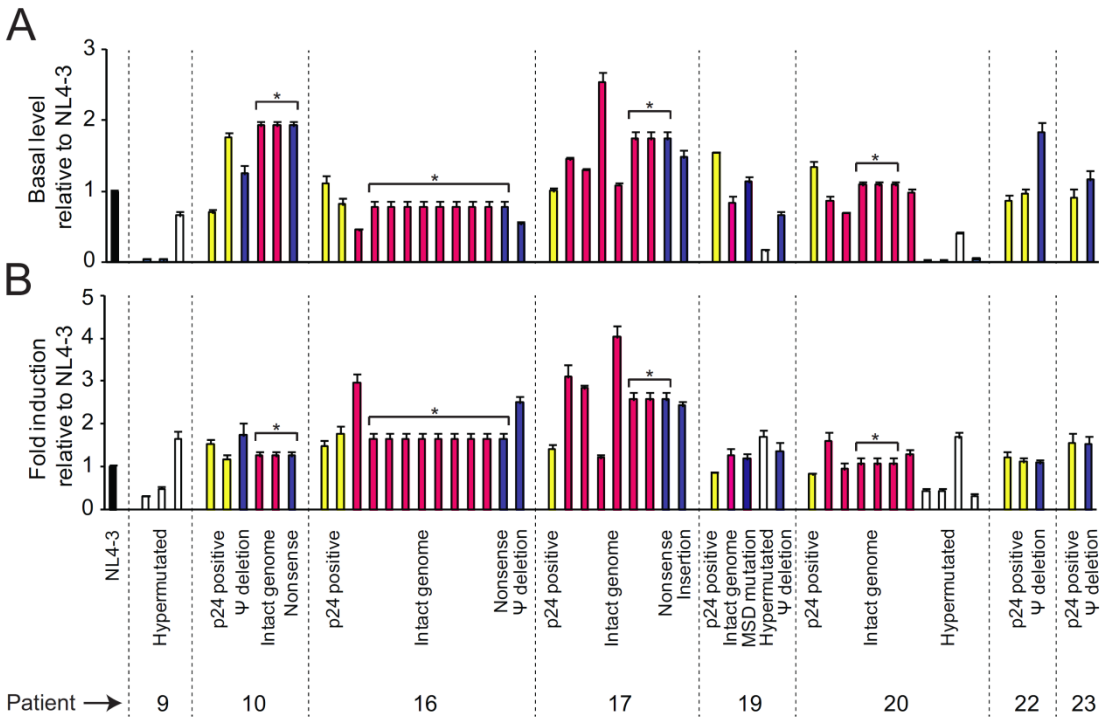

Patient 17
HXB2 coordinates:
140
564
672
U3
R
U5
A
B
U5
17CB2_C1
17CB6_p24_positive
17TC1_LTR_9-3LTRIn
17CB7_3G7
17TC1_LTR_2-3LTRIn
17TC1_LTR_4-3LTRIn
17TC1_LTR_6-3LTRIn
17TC1_LTR_3-3LTRIn
17TC1_LTR_8-3LTRIn
17TC1_LTR_5-3LTRIn
17TC1_LTR_10-3LTRIn
NL4-3
CGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGC

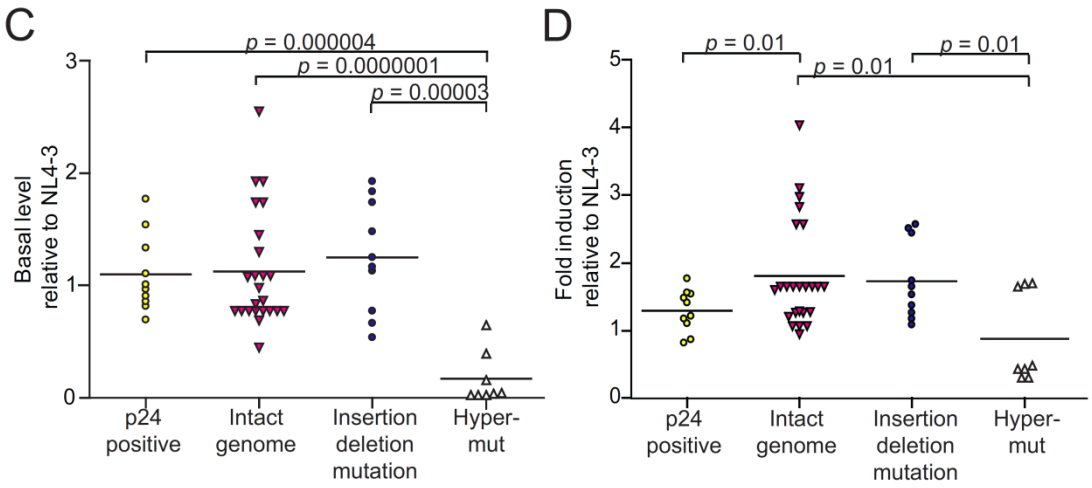

Patient 20
HXB2 coordinates:
140
564
672
U3
R
U5
A
B
U5
20TB13_LTR_22-3LTRIn
20TB13_LTR_5-3LTRIn
20TB13_LTR_16-3LTRIn
20TB13_LTR_23-3LTRIn
20TB13_LTR_11-3LTRIn
20TB13_LTR_27-3LTRIn
20TB13_LTR_31-3LTRIn
20TB13_LTR_18-3LTRIn
20TB13_LTR_6-3LTRIn
20TB13_LTR_53_2-3LTRIn
20TB13_LTR_53_3-3LTRIn
20TB13_LTR_53_4-3LTRIn
20TB13_LTR_30-3LTRIn
20TB13_LTR_28-3LTRIn
20TB13_LTR_19-3LTRIn
20TB13_LTR_29-3LTRIn
20TB13_LTR_10-3LTRIn
20TB13_LTR_53_1-3LTRIn
20CB3_p24_positive
20CB7_47E12
20CB13_LTR_13-3LTRIn
20TB13_LTR_13-3LTRIn
20TB13_LTR_25-3LTRIn
20TB13_LTR_14-3LTRIn
20CB3_LTR_53_3-3LTRIn
20CB13_LTR_16-3LTRIn
20TB13_LTR_7-3LTRIn
20CB13_LTR_25-3LTRIn
20CB13_LTR_12-3LTRIn
20CB3_LTR_53_1-3LTRIn
20CB13_LTR_7-3LTRIn
20TB13_LTR_12-3LTRIn
20TB13_LTR_21-3LTRIn
20CB13_LTR_19-3LTRIn
20CB13_LTR_5-3LTRIn
20CB13_LTR_17-3LTRIn
20CB13_LTR_23-3LTRIn
20CB13_LTR_11-3LTRIn
20CB13_LTR_14-3LTRIn
20CB13_LTR_6-3LTRIn
20CB13_LTR_9-3LTRIn
20CB13_LTR_18-3LTRIn
20TB13_LTR_8-3LTRIn
20CB13_LTR_4-3LTRIn
20TB13_LTR_24-3LTRIn
20TB13_LTR_20-3LTRIn
20TB13_LTR_26-3LTRIn
20CB13_LTR_21-3LTRIn
20CB13_LTR_22-3LTRIn
NL4-3
CGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGC

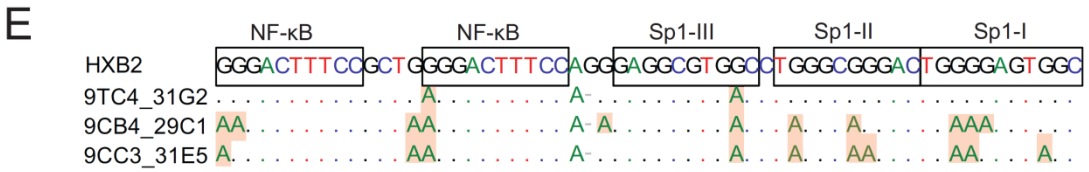
Figure 8. LTR activity of non-induced proviruses

LTR activity measured before (A, C) and four hrs after PMA/ionomycin stimulation (B, D). Resting $CD4^+$ T cells were transfected with LTR-firefly luciferase reporter constructs containing LTRs from non-induced proviruses or induced proviruses from p24 positive wells (yellow bars). Firefly luciferase activity was normalized to an internal control, *Renilla* luciferase driven by a constitutive thymidine kinase promoter. The resulting values were then expressed relative to the NL4-3 control. Non-induced proviruses included those with intact genomes (red bars), insertions or deletions (blue bars), and hypermutation (white bars). Asterisks indicate different clones with the same LTR sequence.

(E) Impact of G→A hypermutation on transcription factor binding sites. Hypermutated sites in representative clones are shaded.


A
Basal level relative to NL4-3
B
Fold induction relative to NL4-3
NL4-3
Hypermutated
p24 positive Ψ deletion
Intact genome Nonsense
p24 positive
Intact genome
Nonsense Ψ deletion
p24 positive
Intact genome
Nonsense Insertion
p24 positive Intact genome MSD mutation Hypermutated Ψ deletion
p24 positive
Intact genome
Hypermutated
p24 positive Ψ deletion
p24 positive Ψ deletion
Patient →
9
10
16
17
19
20
22
23


C
p = 0.000004
p = 0.0000001
p = 0.00003
Basal level relative to NL4-3
p24 positive
Intact genome
Insertion deletion mutation
Hyper-mut
D
p = 0.01
p = 0.01
p = 0.01
Fold induction relative to NL4-3
p24 positive
Intact genome
Insertion deletion mutation
Hyper-mut


E
NF-κB
NF-κB
Sp1-III
Sp1-II
Sp1-I
HXB2 GGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGC
9TC4_31G2
9CB4_29C1
9CC3_31E5

Chapter 5. Non-induced proviruses lack epigenetic silencing markers

Introduction

The presence of intact genome and replication competence indicates intact non-induced proviruses contain required viral components to become infectious once reactivated. We then explored whether these non-induced proviruses are within favorable cellular conditions for reactivation. First, the compact heterochromatin needs to be released into relaxed euchromatin for transcription factors to access binding sites (Tamaru, 2010). Although several groups reported HIV-1 latency in cell-line models is associated with integration into heterochromatin and alphoids repeats (Blazkova et al., 2009; Jordan et al., 2003; Verdin et al., 1993), latent clone selection process in cell line models may represent a biased population of integration into heterochromatin and may not reflect latency conditions *in vivo*. Integration site analysis in patient cells suggests integration into active transcription units *in vivo* (Han et al., 2004). Whether non-induced proviruses are a biased population which are integrated into heterochromatin needs to be investigated to understand whether they could potentially be reactivated.

Second, local chromatin structure is essential for transcriptionally permissive state (Deaton and Bird, 2011). Unmethylated CpG at promoter region is associated with nucleosome displacement and constitutive RNA polymerase II binding, allowing transcription factors to bind (Deaton and Bird, 2011). Methylated promoter CpG inhibits transcription factor binding or recruits methyl-binding domain proteins, which further recruit corepressor complexes and promotes histone deacetylation and therefore transcriptional repression (Deaton and Bird, 2011). Although HIV-1 promoter is

associated with heavy CpG methylation in some cell-line models (Blazkova et al., 2009; Kauder et al., 2009), analysis of HIV-1 promoter CpG methylation led to contradictory results in patient samples (Blazkova et al., 2009; Blazkova et al., 2012b). Here we analyzed CpG methylation of HIV-1 promoter in non-induced proviruses to understand whether non-induced proviruses have favorable or unfavorable local chromatin environment for transcription and reactivation.

Methods

Integration site and CpG methylation

Integration site analysis was carried out by inverse PCR as previously described (Han et al., 2004). Transcription levels of the host genes containing integration sites were determined using a previously described serial analysis of gene expression (SAGE) analysis of Bcl-2-transduced primary resting and activated $CD4^+$ T cells (Shan et al., 2011). CpG methylation analysis was carried out at limiting dilution by bisulfite sequencing as described by Blazkova et al. (Blazkova et al., 2012b).

Results

Most non-induced proviruses are integrated into active transcription units

We determined the locations of non-induced proviruses in the host genome to understand whether they were integrated into sites unfavorable for transcription. Pioneering studies by Bushman *et al.* showed that HIV-1 typically integrates into transcription units (Schroder et al., 2002). However, in some model systems, integration into regions of heterochromatin is associated with latency (Jordan et al., 2003). Using

inverse PCR at limiting dilution, we found that 92.9% of 70 non-induced proviruses resided in transcription units (Figure 9A), consistent with previous observations in patient resting $CD4^+$ T cells (91%) (Han et al., 2004). Using transcript levels measured in a SAGE library from a primary cell model of HIV-1 latency (Shan et al., 2011), we found that most non-induced proviruses were integrated into genes that were transcribed at moderate levels in both resting and activated $CD4^+$ T cells (Figure 9B). Non-induced proviruses were found in both orientations with respect to the host genes (Figure 9C). Overall these results indicate that non-induced proviruses are not integrated into chromosomal regions that are repressive for transcription, and thus other factors must have prevented expression of these proviruses.

Lack of CpG methylation in the LTR of non-induced proviruses

We next examined whether non-induced proviruses were epigenetically silenced. Unlike histone modifications, which can only be analyzed for a population, promoter CpG methylation status can be analyzed at the single provirus level. CpG islands are present in the HIV-1 genome (Chavez et al., 2011), including one in the U3 region of the LTR which contains critical transcription factor binding sites (Figure 10). CpG methylation in the LTR could lead to recruitment of methyl-CpG binding proteins which mediate transcriptional silencing by recruiting nucleosome remodeling and histone deacetylation complexes (Kauder et al., 2009). Therefore, DNA from freshly isolated resting $CD4^+$ T cells and from cells in p24 negative wells was treated with bisulfite, and then the LTR region was amplified under limiting dilution conditions (Blazkova et al., 2012b). Direct sequencing was used to determine the extent of CpG methylation. Only

3.1% of the LTR CpGs were methylated in resting $CD4^+$ T cells from study patients. Even fewer (0.9%) of the CpGs in the LTRs of non-induced proviruses were methylated (Figure 10). In contrast, we readily detected methylation at a CpG island in the *env* region (75.5%), indicating that this method did not selectively amplify non-methylated sites. Although CpG methylation at the LTR clearly is well documented in some models of HIV-1 latency (Kauder et al., 2009), our results indicate that non-induced proviruses are not silenced through CpG methylation at the 5' LTR.

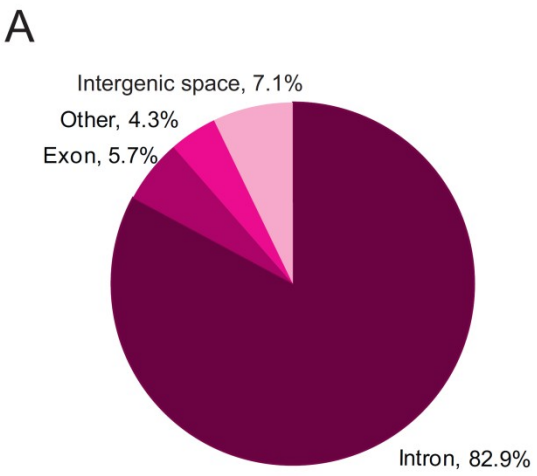
Discussion

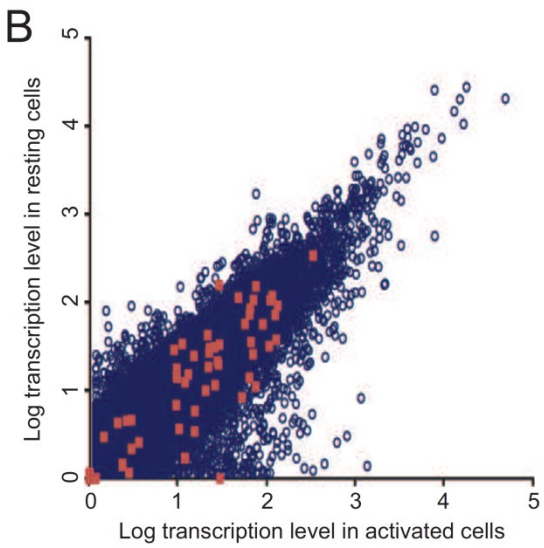
Understanding why intact non-induced proviruses did not produce infectious virus after maximum *in vitro* T cell activation is critical for determining their clinical significance. Possible explanations include silencing by repressive chromatin modifications or transcriptional interference. We analyzed CpG methylation of the LTRs of non-induced proviruses at the clonal level. In contrast to some *in vitro* models of HIV-1 latency (Kauder et al., 2009), we found that in patient $CD4^+$ T cells, there was little CpG methylation at the LTR, consistent with another recent study (Blazkova et al., 2012b). We also examined whether non-induced proviruses are silenced by integration into heterochromatin. We found that most of the non-induced proviruses were integrated into active transcription units, consistent with previous studies showing that most HIV-1 proviruses are integrated into introns of actively transcribed genes in cell lines (Schroder et al., 2002) and patient resting $CD4^+$ T cells (Han et al., 2004). Another potential explanation for the non-induced proviruses is transcriptional interference (Han et al., 2008; Lenasi et al., 2008). Since T cell activation may overcome transcriptional

interference due to the high affinity of NF-κB for its binding sites in the LTR (Lenasi et al., 2008), transcriptional interference may not be a major cause of silencing of the non-induced proviruses.

Figure 9. Integration sites of non-induced proviruses

(A) Distribution of integration sites of non-induced proviruses. Intron and exon boundaries determined as previously described (Shan et al., 2011). Refseq, EST, Human mRNA and Unigene databases were used for intron and exon boundary determination. Other, integration sites matched with other gene or transcription databases but not Refseq. (B) Transcription level of host genes in which non-induced proviruses were integrated. Transcription levels were determined by SAGE analysis of Bcl-2 transduced activated and resting primary $CD4^+$ T cells as previously described (Shan et al., 2011). Blue circles, transcript levels in resting and activated $CD4^+$ T cells. Red squares, transcript levels for the subset of genes in which integration sites of non-induced proviruses were mapped. (C) Transcriptional orientation of non-induced proviruses relative to the host gene.


A
Intergenic space, 7.1%
Other, 4.3%
Exon, 5.7%
Intron, 82.9%


B
Log transcription level in resting cells
Log transcription level in activated cells
0
1
2
3
4
5

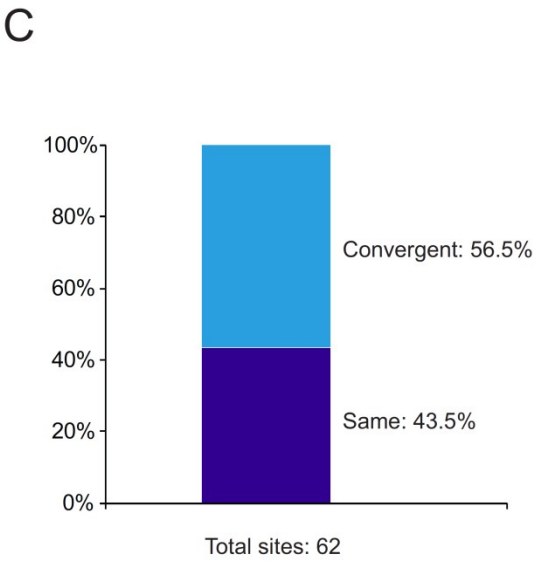

C
100%
80%
60%
40%
20%
0%
Convergent: 56.5%
Same: 43.5%
Total sites: 62

Figure 10. CpG methylation of non-induced proviruses

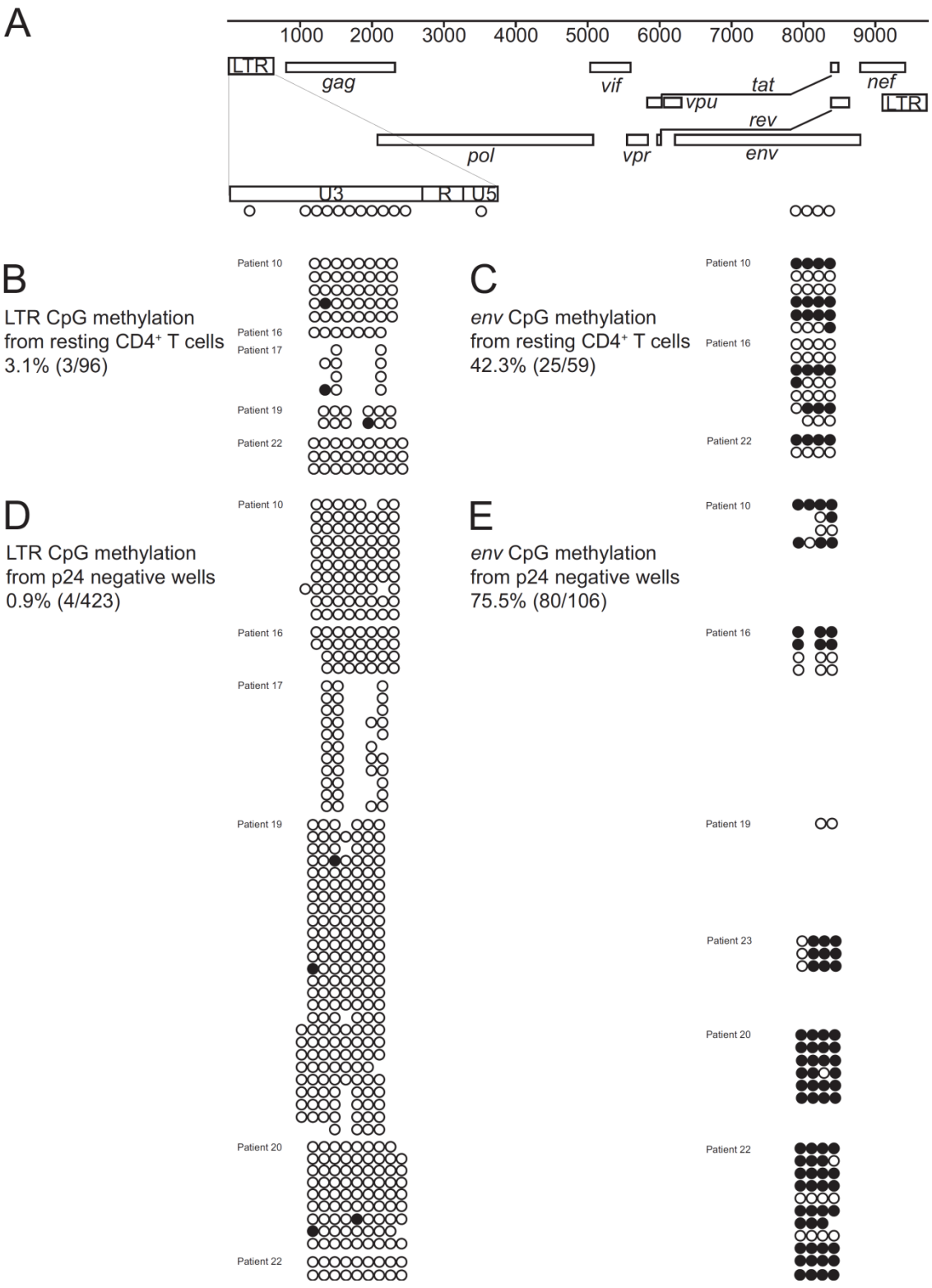
(A) Positions of CpGs analyzed.

(B) CpG methylation in the 5' LTR of proviruses in resting $CD4^+$ T cells from the indicated patients. Each row represents a single provirus amplified by limiting dilution PCR after bisulfite treatment. Open circles, nonmethylated CpGs. Closed circles, methylated CpGs. Missing circles, no CpG present due to sequence polymorphism.

(C) CpG methylation in the *env* gene of proviruses from resting $CD4^+$ T cells.

(D) CpG methylation in the 5' LTR of non-induced proviruses from p24 negative wells.

(E) CpG methylation in the *env* gene of non-induced proviruses from p24 negative wells.


A
1000
2000
3000
4000
5000
6000
7000
8000
9000
LTR
gag
vif
tat
nef
vpu
LTR
rev
pol
vpr
env
U3
R
U5
B
LTR CpG methylation
from resting CD4+ T cells
3.1% (3/96)
Patient 10
Patient 16
Patient 17
Patient 19
Patient 22
C
env CpG methylation
from resting CD4+ T cells
42.3% (25/59)
Patient 10
Patient 16
Patient 22
D
LTR CpG methylation
from p24 negative wells
0.9% (4/423)
Patient 10
Patient 16
Patient 17
Patient 19
Patient 20
Patient 22
E
env CpG methylation
from p24 negative wells
75.5% (80/106)
Patient 10
Patient 16
Patient 19
Patient 23
Patient 20
Patient 22

Chapter 6. Intact non-induced proviruses may increase latent reservoir size by ~60 fold

Introduction

Our discovery of intact non-induced proviruses implies potential underestimation of the HIV-1 latent reservoir. The viral outgrowth assay has been used as the standard measurement of the size of the latent reservoir, although all the latency reversing agents to date can only reactivate a much smaller fraction of the latent reservoir, compared with that of the single round of PHA activation used in viral outgrowth assay (Blazkova et al., 2012a). We demonstrated that intact non-induced proviruses may be an underestimated clinical threat in addition to the inducible reservoir by PHA activation. Identification of this increased barrier challenges the shock-and-kill HIV-1 eradication strategy.

Characterization of non-induced proviruses provides the molecular basis of HIV-1 measurement. Proviral DNA measurements have been used to measure the size of the latent reservoir. However, we found that 88% of the proviruses are defective. Measuring and monitoring defective proviruses does not reflect the true size of the clinically important latent reservoirs. Rather, using PCR-based proviral DNA measurements to monitor the size of the latent reservoir may overestimate, leading to unnecessary and prolonged exposure to latency reversing agents in patients.

Various assays had been used to measure the size of the latent reservoir with highly diverse results (Eriksson et al., 2013). We demonstrated here the individual variability and comparison between different estimates of the most reasonable measurement of the latent reservoir.

Methods

Quantitative DNA PCR

Total proviral DNA in resting $CD4^+$ T cells was quantified using a real-time PCR assay (Durand et al., 2012; Palmer et al., 2003). Copy numbers of the human RNaseP gene copies were measured in separate reactions to quantify cell number (TaqMan Copy Number Reference Assay, RNaseP, Invitrogen) (Spivak et al., 2011).

Statistical analysis

The correlation between viral outgrowth and intact non-induced provirus frequency was analyzed using MedCalc software. Bayesian inference was used to estimate the proportion of intact non-induced genomes in each patient. An empirical Bayesian prior, among five models (Table 3), was chosen to reflect the distribution of the observed data. These proportions were multiplied by total HIV-1 DNA copies per 10^6 resting $CD4^+$ T cells to yield an estimate of intact non-induced proviruses per million cells.

Estimate the frequency of intact non-induced proviruses

The number of intact proviral genomes per million cells was estimated for each patient by Bayesian inference, in collaboration with Sarah B. Laskey and Dr. Daniel I. S. Rosenbloom. The clones examined were modeled as binomial data, with an unknown probability of being intact. The probability was estimated for each patient based on five different prior distributions, in order to emphasize the effect of the choice of prior

distribution on results. The beta distribution is the conjugate prior for a binomial distribution and describes the probabilities of failure and success for a binomial experiment, *i.e.*, the probability of an examined clone being intact versus defective, so all five priors are beta distributions with varying parameters.

The uniform and Jeffreys prior distributions are uninformative priors, chosen to minimize the effect of initial assumptions (Brown et al., 2001). The empirical Bayesian prior distribution was determined by calculating the sample mean and variance of the observed proportions of intact non-induced clones and calculating the beta distribution that fits those parameters. The weighted empirical Bayesian prior has a weighted mean and sample variance, using the number of clones tested for each patient as the weight. The weighting decreases the sample variance and makes the prior stronger. By using the weighting, we acknowledge the consequently strong prior and claim that if we were to double or triple the number of clones tested, we would probably find at least one intact non-induced provirus in every patient. The three patients for whom no intact clones were found are also the three patients in whom the fewest clones were tested. The iterative median approach started with an arbitrary prior, used the posterior median to create an empirical Bayes prior for the next round, and iterated until convergence. (Using the empirical Bayes prior as the initial prior, 40 iterations were sufficient for convergence.)

The empirical Bayesian priors reflect the distribution of the observed data, so these are not fully Bayesian analyses. However, empirical Bayesian analysis can be understood as an approximation of the fully Bayesian hierarchical approach (Gelman et al., 2004).

The median and 95% credible intervals for each approach are reported in Table 3. To estimate the number of cells harboring intact non-induced proviruses per million resting $CD4^+$ T cells, the median and 95% interval were multiplied by the number of cells carrying HIV-1 DNA per million resting $CD4^+$ T cells.

Results

Although most non-induced proviruses have identifiable lethal defects, a substantial fraction are intact and replication-competent at the primary sequence level. Analysis of LTR function, integration sites, and methylation status suggests that these intact non-induced proviruses could be induced *in vivo*, thereby increasing latent reservoir size. We compared the frequency of induced proviruses (defined using the viral outgrowth assay) and intact non-induced proviruses (quantitated as the product of total proviral DNA frequency and the fraction of non-induced proviruses that are intact) among the total pool of proviruses (measured by quantitative real-time PCR). Bayesian analysis was chosen instead of maximum likelihood estimation because the former provides nonzero point estimates for patients from whom no clones with intact genomes were identified. The positive viral outgrowth assay results in every patient and the successful detection of intact non-induced proviruses in patients for whom >20 clones were analyzed suggest that intact non-induced proviruses could be detected in every patient if enough clones are examined. The fraction of intact non-induced proviruses was calculated as the median of an empirical Bayesian posterior, the most conservative of five models tested (Table 3), with a prior distribution chosen to reflect the observed data. Both the fraction of intact non-induced proviruses and the total number of proviruses per

million resting $CD4^+$ T cells varied dramatically from patient to patient (Figure 11A). There was no correlation between the viral outgrowth assay and the frequency of intact non-induced proviruses (Figure 11B). All statistical models (Table 3) indicated that the median frequency of intact non-induced proviruses was at least ~60 fold higher than the frequency of induced proviruses detected in the viral outgrowth assay. If the intact non-induced proviruses described here can be induced *in vivo*, then the size of the latent reservoir is much greater than previously thought (Figure 11C).

Discussion

To prove replication-competence, we reconstructed six intact, non-induced proviruses from six different p24 negative wells from four patients. Surprisingly, all reconstructed viruses replicated as well as the standard reference isolate and control viruses reconstructed from p24 positive wells. A sterilizing cure requires elimination of all replication-competent HIV-1, and therefore the discovery that intact non-induced proviruses are replication-competent means that the number of proviruses that must be eliminated is much higher than previously thought. We conservatively estimate that the number may be ~60 fold higher than estimates based on the viral outgrowth assay. Some statistical models suggest an even higher number (medians of 97 – 273 fold). Importantly, there is large inter-patient variation in this and other measures of latent reservoir size. Overall, our results indicate that the "shock and kill" strategy (Archin et al., 2012; Deeks, 2012) is challenged with a large but unmeasured hidden population of replication-competent proviruses. Interestingly, despite the intense search for novel latency reversing agents, none of the drugs tested to date reaches the robust level of *in*

vitro HIV-1 induction achieved by PHA. Thus the finding that the true size of the latent reservoir may be ~60 fold greater than that estimated using PHA activation is particularly disturbing. However, it is also important to point out that the even a low level of virus gene expression may be sufficient to allow the elimination of infected cells by an appropriately primed CTL response (Shan et al., 2012), and that the critical variable may the fraction of latently infected cells induced to express HIV-1 genes.

Table 3. Estimation of intact non-induced provirus frequency

Patient ID #	Viral outgrowth (Infectious units per million resting $CD4^+$ T cells)	Total HIV-1 DNA (Copies per million resting $CD4^+$ T cells)	Intact non – induced proviral frequency[a,b]					
			Observed	Empirical Bayes posterior estimate	Jeffreys posterior estimate	Weighted empirical Bayes posterior estimate	Median iteration posterior estimate	Uniform posterior estimate
9	0.4169	1,431	0 0.0% (0/16)	13 (0–164) 0.9% (0.0%–11.4%) 32X (0X–392X)	20 (0–205) 1.4% (0.0%–14.3%) 48X (0X–491X)	24 (0–186) 1.7% (0.0%–13.0%) 59X (0–447X)	30 (1–187) 2.1% (0.0%–13.1%) 72X (1X–448X)	57 (2–279) 4.0% (0.1%–19.5%) 137X (5X–669X)
10	0.5109	10,117	430 4.3% (2/47)	425 (81–1,215) 4.2% (0.8%–12.0%) 832X (158X–2,378X)	465 (91–1,311) 4.6% (0.9%–13.0%) 910X (177X–2,566)	453 (94–1,245) 4.5% (0.9%–12.3%) 886 (184–2,437)	459 (100–1,233) 4.5% (1.0%–12.2%) 898 (196–2,412)	560 (132–1,442) 5.5% (1.3%–14.3%) 1,095X (259X–2,823X)
16	8.6505	205	49 23.7% (9/38)	45 (23–74) 21.9% (11.3%–35.9%) 5X (3X–9X)	49 (25–80) 23.9% (12.4%–38.8%) 6X (3X–9X)	45 (23–73) 21.7% (11.3%–35.5%) 5X (3X–8X)	43 (22–70) 20.9% (10.9%–34.2%) 5X (3X–8X)	50 (27–81) 24.6% (13.0%–39.3%) 6X (3X–9X)
17	2.7659	17,377	6,082 35.0% (7/20)	5,247 (2,511–8,693) 30.2% (14.4%–50.0%) 1,897X (908X–3,143X)	6,126 (2,994–9,866) 35.3% (17.2%–56.8%) 2,215X (1,082X–3,567X)	5,118 (2,471–8,466) 29.5% (14.2%–48.7%) 1,851X (893X–3,061X)	4,769 (2,305–7,939) 27.4% (13.3%–45.7%) 1,724X (833X–2,870X)	6,246 (3,147–9,900) 35.9% (18.1%–57.0%) 2,258X (1,138X–3,579X)
19	1.1076	334	10 3.1% (1/32)	11 (1–41) 3.2% (0.3%–12.3%) 10X (1X–37X)	12 (1–46) 3.7% (0.3%–13.7%) 11X (1X–41X)	12 (1–43) 3.6% (0.4%–12.8%) 11X (1X–38X)	12 (2–42) 3.7% (0.5%–12.6%) 11X (1X–38X)	17 (2–53) 5.0% (0.7%–15.8%) 15X (2X–47X)
20	1.6097	44,213	8,290 18.8% (6/32)	7,605 (3,245–13,972) 17.2% (7.3%–31.6%) 4,725X (2,016X–8,680X)	8,436 (3,637–15,311) 19.1% (8.2%–34.6%) 5,241X (2,260X–9,512X)	7,580 (3,284–13,821) 17.1% (7.4%–31.3%) 4,709X (2,040X–8,586X)	7,285 (3,182–13,259) 16.5% (7.2%–30.0%) 4,526X (1,977X–8,237X)	8,846 (3,971–15,678) 20.0% (9.0%–35.5%) 5,495X (2,467X–9,740X)
22	1.6253	314	0 0.0% (0/18)	3 (0–33) 0.9% (0.0%–10.4%) 2X (0X–20X)	4 (0–40) 1.2% (0.0%–12.9%) 2X (0X–25X)	5 (0–38) 1.6% (0.0%–12.0%) 3X (0X–23X)	6 (0–38) 1.9% (0.0%–12.1%) 4X (0X–23X)	11 (0–55) 3.6% (0.1%–17.6%) 7X (0X–34X)
23	0.4169	2,794	0 0.0% (0/10)	38 (0–449) 1.4% (0.0%–16.1%) 91X (0X–1,078X)	61 (0–607) 2.2% (0.0%–21.7%) 147X (0X–1,456X)	67 (1–499) 2.4% (0.0%–17.9%) 161X (1X–1,197X)	79 (1–482) 2.8% (0.0%–17.3%) 189X (3X–1,157X)	171 (6–796) 6.1% (0.2%–28.5%) 409X (15X–1,910X)
Median	1.35865	2,112	30 3.7% (25/213)	41 (12–306) 3.7% (0.5%–14.2%) 62X (2X–735X)	55 (13–406) 4.1% (0.6%–18.0%) 97X (2X–974X)	56 (12–343) 4.0% (0.7%–15.4%) 110X (2X–822X)	61 (12–335) 4.1% (0.7%–15.2%) 130X (3X–803X)	114 (17–538) 5.8% (1.0%–24.0%) 273X (10X–1,289X)

[a]Numbers, median of intact non-induced provirus copy numbers per million resting $CD4^+$ T cells (95% credible interval). %, percentage of intact non-induced proviruses (95% credible interval). X: fold difference between intact non-induced proviruses and viral outgrowth (95% credible interval).

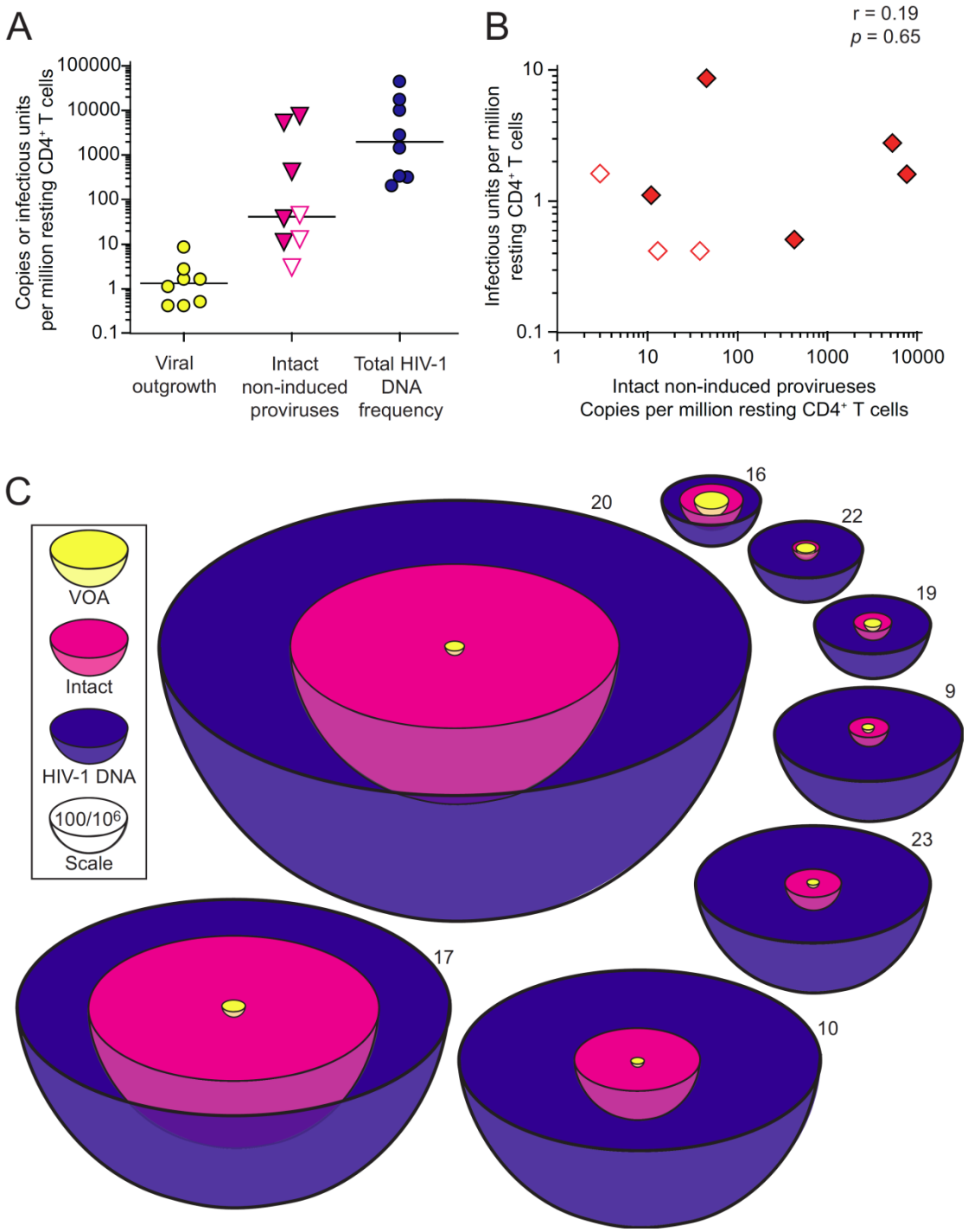
[b]Bayesian analysis was chosen instead of maximum likelihood estimation because the former provides nonzero point estimates for patients from whom no clones with intact genomes were identified. The positive VOA results in every patient and the successful detection of intact non-induced proviruses in patients for whom >20 clones were analyzed suggest that intact non-induced proviruses could be detected in every patient if enough clones are examined.

Figure 11. Quantification of intact non-induced proviruses

(A) Comparison of the frequency of cells detected in the viral outgrowth assay, cells carrying intact non-induced proviruses, and cells carrying HIV-1 DNA. The frequency of cells with HIV-1 DNA was measured by quantitative PCR on freshly isolated resting $CD4^+$ T cells. The frequency of cells with intact, non-induced proviruses was calculated as frequency of cells with HIV-1 DNA times the proportion of intact non-induced proviruses estimated for each patient using an empirical Bayesian model. Open symbols, no intact proviruses detected, empirical Bayesian estimate plotted. Bars, median value.

(B) Correlation between the frequency of cells detected in the viral outgrowth assay and the frequency of cells with intact non-induced proviruses. Open symbols, see (A).

(C) Scale representation of the frequencies of the infected resting $CD4^+$ T cell populations. Volume reflects population size. Yellow circles, minimum size of the latent reservoir as measured by viral outgrowth assay. Magenta, frequency of cells with intact proviruses, calculated as the frequency of cells detected in the viral outgrowth assay plus the frequency of cells with intact non-induced proviruses. This is the potential size of the latent reservoir if intact non-induced proviruses can be induced *in vivo*. Blue, cells with HIV-1 DNA.


A
Copies or infectious units per million resting CD4+ T cells
100000
10000
1000
100
10
1
0.1
Viral outgrowth
Intact non-induced proviruses
Total HIV-1 DNA frequency
B
r = 0.19
p = 0.65
Infectious units per million resting CD4+ T cells
10
1
0.1
1
10
100
1000
10000
Intact non-induced provirueses Copies per million resting CD4+ T cells
C
VOA
Intact
HIV-1 DNA
100/10^6
Scale
20
16
22
19
9
23
17
10

Chapter 7. Stochastic activation of non-induced proviruses

Introduction

Although non-induced proviruses were not reactivated after one round of maximum T cell activation *in vitro*, it does not mean that it will never be induced *in vivo*. Whether and why non-induced proviruses can be reactivated *in vivo* is the key question to HIV-1 eradication.

It has been believed that maximum T cell activation leads to maximum HIV-1 activation. Theoretically, our demonstration of 100% T cell activation by PHA should lead to 100% HIV-1 provirus reactivation, and therefore none of the non-induced proviruses can be reactivated. However, HIV-1 activation may not follow T cell activation mechanisms. Variable levels of Tat, which is required to form positive transcriptional elongation factor b (pTEFb) to released stalled RNA polymerase II from LTR, may lead to stochastic fluctuations and phenotypic variability(Weinberger et al., 2005). Viral fate determination has been proposed to be determined by transcriptional bursting of Tat (Singh et al., 2010; Weinberger et al., 2008). If stochasticity is the major mechanism of HIV-1 reactivation, it means that every single non-induced proviruses might become activated due to stochasticity, and cure cannot be achieved before eradicating all the intact non-induced proviruses.

To determine whether these non-induced proviruses can be reactivated by stochasticity, we examined whether repeated PHA activation can activate additional non-induced proviruses.

Methods

Resting $CD4^+$ T cells from four aviremic donors were plated at 0.2 million/well and subjected to activation with PHA and irradiated allogeneic PBMC for one day. Patient cells were then cultured in a transwell viral outgrowth assay, with patient cells in the upper chamber along with irradiated allogeneic PBMC and donor lymphoblasts in the lower chamber. On day seven, viability of cells in the upper chamber was examined by forward and side scattered flow cytometry. On day seven of culture, the cells in the upper and lower chambers of each well were combined and then separated equally into duplicate culture wells. As all patient cells had divided at least once by day seven, each pair of split wells should contain an approximately equal number of daughter cells derived from the patient cells in the original well. One set of the split culture wells was subjected to an additional round of stimulation with PHA and irradiated allogeneic PBMC. The other set of wells were not stimulated. Both sets of wells were cultured for an additional 14 days with two more additions of donor lymphoblasts. Levels of HIV-1 p24 antigen were measured in the supernatant on day 21.

Results

To determine whether intact non-induced proviruses are permanently silenced or potentially inducible under certain conditions, we tested whether repeated PHA stimulation can induce additional non-induced proviruses (Figure 12A). We stimulated multiple replicate cultures of 2×10^5 patient resting $CD4^+$ T cells with PHA in viral outgrowth assay conditions. We then split each culture well equally into two wells on day seven. As all patient cells have divided by day seven (Figure 1), each split well

contained daughter cells derived from cells activated in the original well. One set of the split cultures wells were activated again with PHA while the other set was cultured without additional stimulation. We then compared supernatant p24 levels after another 14 days of culture. Among 126 p24 negative wells from four patients, 31 (24.6%) became p24 positive after the additional round of PHA stimulation, while their paired culture well that did not receive an additional round of stimulation remained p24 negative (Figure 12B). These results demonstrate that at least some intact non-induced proviruses can be induced under repeated stimulation.

Discussion

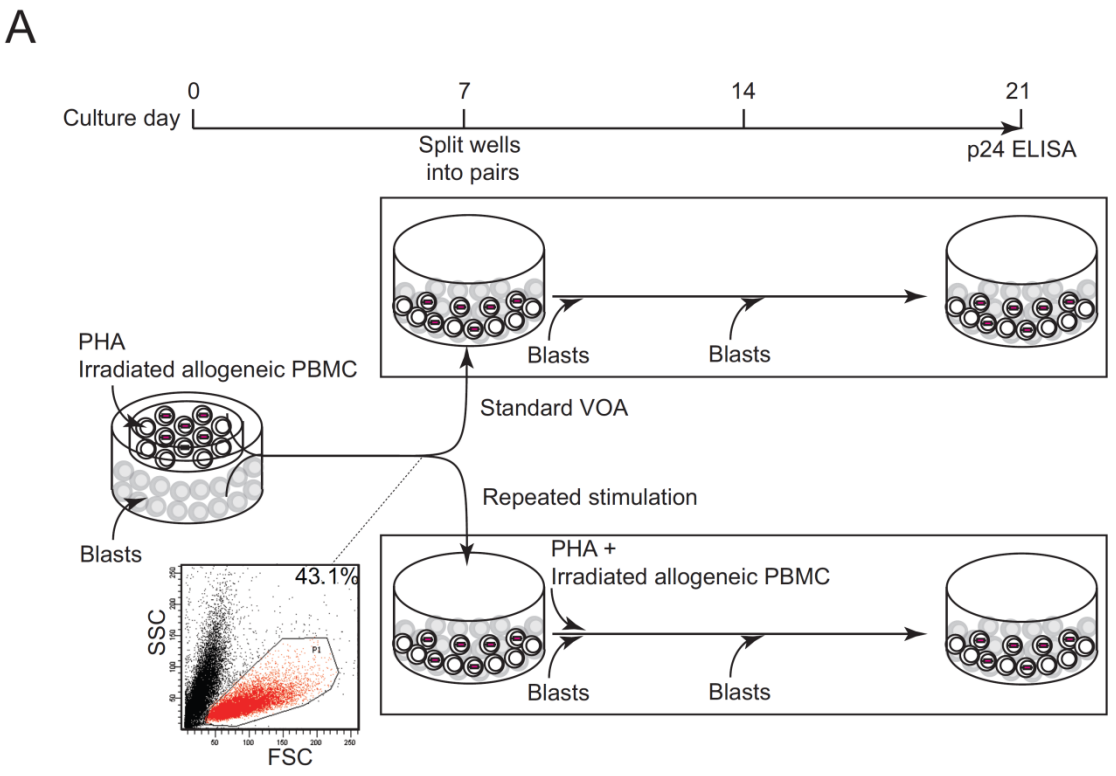
We propose that despite maximum T cell activation, the induction of latent proviruses is stochastic. Cellular gene expression levels may follow a digital or analogue distribution after T cell receptor activation, as a result of stochastic and dynamic processes (Chakraborty and Das, 2010). Elegant experimental and theoretical studies have shown that HIV-1 proviruses may show stochastic fluctuations in expression depending on levels of Tat (Burnett et al., 2009; Weinberger et al., 2005; Weinberger et al., 2005; Weinberger et al., 2008). We propose that intact proviruses undergo stochastic induction even after maximum cellular activation (Figure 13). Some will be induced by one round of activation, while others will remain silent but retain the potential to be activated subsequently. These results indicate an increased barrier to cure, as all intact non-induced proviruses need to be eradicated. Underestimation of intact proviruses by viral outgrowth assays could be reflected in delayed viral rebound after an apparent "cure", and overestimation of latent reservoir size resulting from detection of defective

proviruses by PCR assays could result in prolonged, excessive exposure to toxic latency reversing agents. Thus, the molecular analysis of non-induced proviruses contributes in an important way to HIV-1 eradication efforts.

Figure 12. Stochastic activation of non-induced proviruses

(A) Scheme for repeated stimulation. To understand whether non-induced proviruses may be induced by repeated stimuli, resting $CD4^+$ T cells from four aviremic donors were plated at 0.2 million/well and subjected to activation with PHA and irradiated allogeneic PBMC for one day. Patients cells were then cultured in a transwell viral outgrowth assay, with patient cells in the upper chamber along with irradiated allogeneic PBMC and donor lymphoblasts in the lower chamber. On day seven, viability of cells in the upper chamber was examined. On day seven of culture, the cells in the upper and lower chambers of each well were combined and then separated equally into duplicate culture wells. As all patient cells had divided at least once by day seven, each pair of split wells should contain an approximately equal number of daughter cells derived from the patient cells in the original well. One set of the split culture wells was subjected to an additional round of stimulation with PHA and irradiated allogeneic PBMC. The other set of wells were not stimulated. Both sets of wells were cultured for an additional 14 days with two more additions of donor lymphoblasts. Levels of HIV-1 p24 antigen were measured in the supernatant on day 21.

(B) Comparison of p24 results for of each pair of standard and repeated stimulation culture wells. Red bars: number of wells containing non-induced proviruses induced to produce replication competent viruses after repeated PHA stimulation. Magenta bars: number of wells containing non-induced proviruses which are not induced after repeated PHA stimulation. Similar results were obtained in experiments in which patient resting $CD4^+$ T cells were plated at one million/well except that the fraction of positive wells was higher.


A
0
7
14
21
Culture day
Split wells into pairs
p24 ELISA
PHA
Irradiated allogeneic PBMC
Blasts
Standard VOA
Repeated stimulation
43.1%
SSC
FSC
P1
PHA +
Irradiated allogeneic PBMC
Blasts
Blasts
Blasts
Blasts

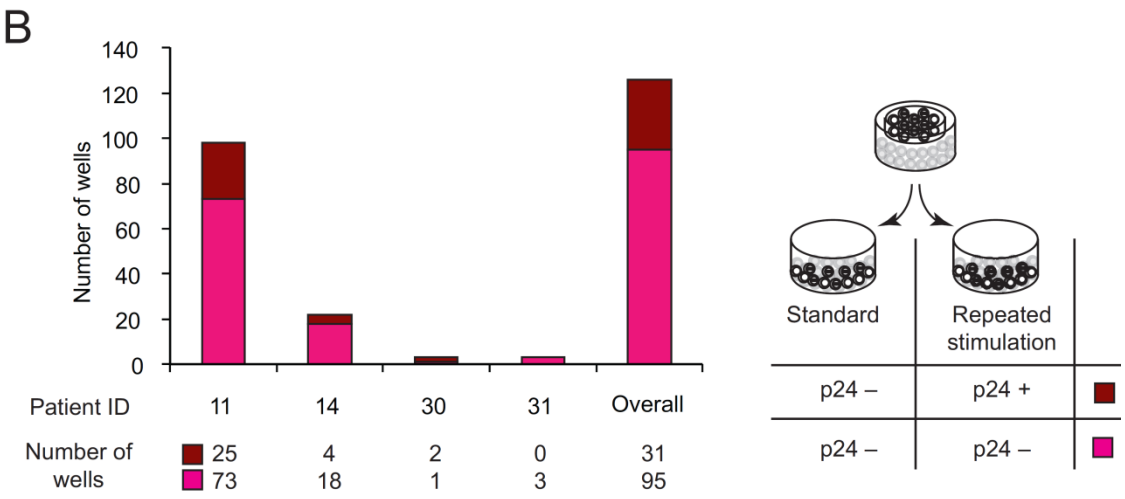
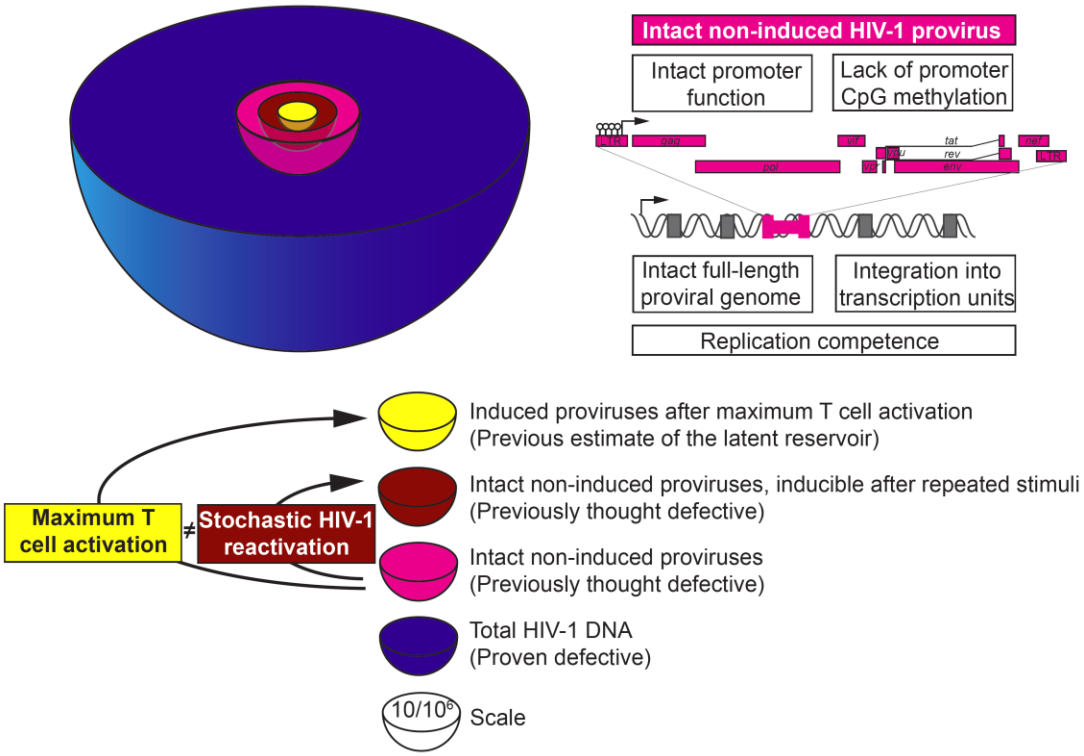

B
Number of wells
140
120
100
80
60
40
20
0
Patient ID
11
14
30
31
Overall
Number of wells
25
73
4
18
2
1
0
3
31
95
Standard
Repeated stimulation
p24 –
p24 +
p24 –
p24 –

Figure 13. Mechanism of HIV-1 latency reactivation



REFERENCES

Archin, N.M., Espeseth, A., Parker, D., Cheema, M., Hazuda, D., and Margolis, D.M. (2009). Expression of latent HIV induced by the potent HDAC inhibitor suberoylanilide hydroxamic acid. AIDS Res. Hum. Retroviruses *25,* 207-212.

Archin, N.M., Liberty, A.L., Kashuba, A.D., Choudhary, S.K., Kuruc, J.D., Crooks, A.M., Parker, D.C., Anderson, E.M., Kearney, M.F., Strain, M.C.*, et al.* (2012). Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. Nature *487,* 482-485.

Barletta, J.M., Edelman, D.C., and Constantine, N.T. (2004). Lowering the detection limits of HIV-1 viral load using real-time immuno-PCR for HIV-1 p24 antigen. Am. J. Clin. Pathol. *122,* 20-27.

Bebenek, K., Abbotts, J., Roberts, J.D., Wilson, S.H., and Kunkel, T.A. (1989). Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J. Biol. Chem. *264,* 16948-16956.

Blankson, J.N., Finzi, D., Pierson, T.C., Sabundayo, B.P., Chadwick, K., Margolick, J.B., Quinn, T.C., and Siliciano, R.F. (2000). Biphasic decay of latently infected CD4+ T cells in acute human immunodeficiency virus type 1 infection. J. Infect. Dis. *182,* 1636-1642.

Blazkova, J., Chun, T.W., Belay, B.W., Murray, D., Justement, J.S., Funk, E.K., Nelson, A., Hallahan, C.W., Moir, S., Wender, P.A., and Fauci, A.S. (2012a). Effect of histone deacetylase inhibitors on HIV production in latently infected, resting CD4(+) T cells from infected individuals receiving effective antiretroviral therapy. J. Infect. Dis. *206,* 765-769.

Blazkova, J., Murray, D., Justement, J.S., Funk, E.K., Nelson, A.K., Moir, S., Chun, T.W., and Fauci, A.S. (2012b). Paucity of HIV DNA methylation in latently infected, resting CD4+ T cells from infected individuals receiving antiretroviral therapy. J. Virol. *86,* 5390-5392.

Blazkova, J., Trejbalova, K., Gondois-Rey, F., Halfon, P., Philibert, P., Guiguen, A., Verdin, E., Olive, D., Van Lint, C., Hejnar, J., and Hirsch, I. (2009). CpG methylation controls reactivation of HIV from latency. PLoS Pathog. *5,* e1000554.

Bohnlein, E., Lowenthal, J.W., Siekevitz, M., Ballard, D.W., Franza, B.R., and Greene, W.C. (1988). The same inducible nuclear proteins regulates mitogen activation of both the interleukin-2 receptor-alpha gene and type 1 HIV. Cell *53,* 827-836.

Brown, L.D., Cai, T.T., and DasGupta, A. (2001). Interval Estimation for a Binomial Proportion. Statistical Science *16(2),* 101.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P., and Schaffer, D.V. (2009). Control of stochastic gene expression by host factors at the HIV promoter. PLoS Pathog. *5,* e1000260.

Chakraborty, A.K., and Das, J. (2010). Pairing computation with experimentation: a powerful coupling for understanding T cell signalling. Nat. Rev. Immunol. *10,* 59-71.

Chavez, L., Kauder, S., and Verdin, E. (2011). In vivo, in vitro, and in silico analysis of methylation of the HIV-1 provirus. Methods *53,* 47-53.

Chun, T.W., Stuyver, L., Mizell, S.B., Ehler, L.A., Mican, J.A., Baseler, M., Lloyd, A.L., Nowak, M.A., and Fauci, A.S. (1997). Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc. Natl. Acad. Sci. U. S. A. *94,* 13193-13197.

Contreras, X., Schweneker, M., Chen, C.S., McCune, J.M., Deeks, S.G., Martin, J., and Peterlin, B.M. (2009). Suberoylanilide hydroxamic acid reactivates HIV from latently infected cells. J. Biol. Chem. *284,* 6782-6789.

Coull, J.J., Romerio, F., Sun, J.M., Volker, J.L., Galvin, K.M., Davie, J.R., Shi, Y., Hansen, U., and Margolis, D.M. (2000). The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. J. Virol. *74,* 6790-6799.

Cujec, T.P., Okamoto, H., Fujinaga, K., Meyer, J., Chamberlin, H., Morgan, D.O., and Peterlin, B.M. (1997). The HIV transactivator TAT binds to the CDK-activating kinase and activates the phosphorylation of the carboxy-terminal domain of RNA polymerase II. Genes Dev. *11,* 2645-2657.

Davey, R.T.,Jr, Bhat, N., Yoder, C., Chun, T.W., Metcalf, J.A., Dewar, R., Natarajan, V., Lempicki, R.A., Adelsberger, J.W., Miller, K.D.*, et al.* (1999). HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. Proc. Natl. Acad. Sci. U. S. A. *96,* 15109-15114.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25,* 1010-1022.

Deeks, S.G. (2012). HIV: Shock and kill. Nature *487,* 439-440.

Depper, J.M., Leonard, W.J., Drogula, C., Kronke, M., Waldmann, T.A., and Greene, W.C. (1985). Interleukin 2 (IL-2) augments transcription of the IL-2 receptor gene. Proc. Natl. Acad. Sci. U. S. A. *82,* 4230-4234.

Duh, E.J., Maury, W.J., Folks, T.M., Fauci, A.S., and Rabson, A.B. (1989). Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction

of nuclear factor binding to the NF-kappa B sites in the long terminal repeat. Proc. Natl. Acad. Sci. U. S. A. *86,* 5974-5978.

Durand, C.M., Ghiaur, G., Siliciano, J.D., Rabi, S.A., Eisele, E.E., Salgado, M., Shan, L., Lai, J.F., Zhang, H., Margolick, J.*, et al.* (2012). HIV-1 DNA is detected in bone marrow populations containing CD4+ T cells but is not found in purified CD34+ hematopoietic progenitor cells in most patients on antiretroviral therapy. J. Infect. Dis. *205,* 1014-1018.

Ehrenberg, P.K., and Michael, N.L. (2005). PCR amplification, cloning, and construction of HIV-1 infectious molecular clones from virtually full-length HIV-1 genomes. Methods Mol. Biol. *304,* 387-398.

Eriksson, S., Graf, E.H., Dahl, V., Strain, M.C., Yukl, S.A., Lysenko, E.S., Bosch, R.J., Lai, J., Chioma, S., Emad, F.*, et al.* (2013). Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. PLoS Pathog. *9,* e1003174.

Finzi, D., Hermankova, M., Pierson, T., Carruth, L.M., Buck, C., Chaisson, R.E., Quinn, T.C., Chadwick, K., Margolick, J., Brookmeyer, R.*, et al.* (1997). Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science *278,* 1295-1300.

Ganesh, L., Burstein, E., Guha-Niyogi, A., Louder, M.K., Mascola, J.R., Klomp, L.W., Wijmenga, C., Duckett, C.S., and Nabel, G.J. (2003). The gene product Murr1 restricts HIV-1 replication in resting CD4+ lymphocytes. Nature *426,* 853-857.

Gao, F., Robertson, D.L., Carruthers, C.D., Morrison, S.G., Jian, B., Chen, Y., Barre-Sinoussi, F., Girard, M., Srinivasan, A., Abimiku, A.G.*, et al.* (1998). A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1. J. Virol. *72,* 5680-5698.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2004). Bayesian Data Analysis, Second Edition.(Boca Raton, FL: Chapman and Hall/CRC).

Han, Y., Lassen, K., Monie, D., Sedaghat, A.R., Shimoji, S., Liu, X., Pierson, T.C., Margolick, J.B., Siliciano, R.F., and Siliciano, J.D. (2004). Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. J. Virol. *78,* 6122-6133.

Han, Y., Lin, Y.B., An, W., Xu, J., Yang, H.C., O'Connell, K., Dordai, D., Boeke, J.D., Siliciano, J.D., and Siliciano, R.F. (2008). Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. Cell. Host Microbe *4,* 134-146.

He, G., and Margolis, D.M. (2002). Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency

virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. Mol. Cell. Biol. *22,* 2965-2973.

Hermankova, M., Siliciano, J.D., Zhou, Y., Monie, D., Chadwick, K., Margolick, J.B., Quinn, T.C., and Siliciano, R.F. (2003). Analysis of human immunodeficiency virus type 1 gene expression in latently infected resting CD4+ T lymphocytes in vivo. J. Virol. *77,* 7383-7392.

Herrmann, C.H., and Rice, A.P. (1995). Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. J. Virol. *69,* 1612-1620.

Jones, K.A., and Peterlin, B.M. (1994). Control of RNA initiation and elongation at the HIV-1 promoter. Annu. Rev. Biochem. *63,* 717-743.

Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J. *22,* 1868-1877.

Kao, S.Y., Calman, A.F., Luciw, P.A., and Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature *330,* 489-493.

Katlama, C., Deeks, S.G., Autran, B., Martinez-Picado, J., van Lunzen, J., Rouzioux, C., Miller, M., Vella, S., Schmitz, J.E., Ahlers, J., Richman, D.D., and Sekaly, R.P. (2013). Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. Lancet *381,* 2109-2117.

Kauder, S.E., Bosque, A., Lindqvist, A., Planelles, V., and Verdin, E. (2009). Epigenetic regulation of HIV-1 latency by cytosine methylation. PLoS Pathog. *5,* e1000495.

Kinoshita, S., Su, L., Amano, M., Timmerman, L.A., Kaneshima, H., and Nolan, G.P. (1997). The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells. Immunity *6,* 235-244.

Laird, G.M., Eisele, E.E., Rabi, S.A., Lai, J., Chioma, S., Blankson, J.N., Siliciano, J.D., and Siliciano, R.F. (2013). Rapid Quantification of the Latent Reservoir for HIV-1 Using a Viral Outgrowth Assay. PLoS Pathog. *9,* e1003398.

Lassen, K., Han, Y., Zhou, Y., Siliciano, J., and Siliciano, R.F. (2004). The multifactorial nature of HIV-1 latency. Trends Mol. Med. *10,* 525-531.

Lehrman, G., Hogue, I.B., Palmer, S., Jennings, C., Spina, C.A., Wiegand, A., Landay, A.L., Coombs, R.W., Richman, D.D., Mellors, J.W.*, et al.* (2005). Depletion of latent HIV-1 infection in vivo: a proof-of-concept study. Lancet *366,* 549-555.

Lenasi, T., Contreras, X., and Peterlin, B.M. (2008). Transcriptional interference antagonizes proviral gene expression to promote HIV latency. Cell. Host Microbe *4,* 123-133.

Li, B., Gladden, A.D., Altfeld, M., Kaldor, J.M., Cooper, D.A., Kelleher, A.D., and Allen, T.M. (2007). Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. J. Virol. *81,* 193-201.

Li, Y., Kappes, J.C., Conway, J.A., Price, R.W., Shaw, G.M., and Hahn, B.H. (1991). Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: identification of replication-competent and -defective viral genomes. J. Virol. *65,* 3973-3985.

Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V., and Cullen, B.R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature *338,* 254-257.

Mansky, L.M., and Temin, H.M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J. Virol. *69,* 5087-5094.

Mustelin, T., Coggeshall, K.M., Isakov, N., and Altman, A. (1990). T cell antigen receptor-mediated activation of phospholipase C requires tyrosine phosphorylation. Science *247,* 1584-1587.

Nabel, G., and Baltimore, D. (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature *326,* 711-713.

Palmer, S., Wiegand, A.P., Maldarelli, F., Bazmi, H., Mican, J.M., Polis, M., Dewar, R.L., Planta, A., Liu, S., Metcalf, J.A., Mellors, J.W., and Coffin, J.M. (2003). New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. J. Clin. Microbiol. *41,* 4531-4536.

Patel, S.S., Duby, A.D., Thiele, D.L., and Lipsky, P.E. (1988). Phenotypic and functional characterization of human T cell clones. J. Immunol. *141,* 3726-3736.

Richman, D.D., Margolis, D.M., Delaney, M., Greene, W.C., Hazuda, D., and Pomerantz, R.J. (2009). The challenge of finding a cure for HIV infection. Science *323,* 1304-1307.

Rose, P.P., and Korber, B.T. (2000). Detecting hypermutations in viral sequences with an emphasis on G --> A hypermutation. Bioinformatics *16,* 400-401.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J.*, et al.* (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463,* 237-240.

Sahu, G.K., Sarria, J.C., and Cloyd, M.W. (2010). Recovery of replication-competent residual HIV-1 from plasma of a patient receiving prolonged, suppressive highly active antiretroviral therapy. J. Virol. *84,* 8348-8352.

Salazar-Gonzalez, J.F., Salazar, M.G., Keele, B.F., Learn, G.H., Giorgi, E.E., Li, H., Decker, J.M., Wang, S., Baalwa, J., Kraus, M.H.*, et al.* (2009). Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. J. Exp. Med. *206,* 1273-1289.

Sanchez, G., Xu, X., Chermann, J.C., and Hirsch, I. (1997). Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. J. Virol. *71,* 2233-2240.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110,* 521-529.

Selby, M.J., and Peterlin, B.M. (1990). Trans-activation by HIV-1 Tat via a heterologous RNA binding protein. Cell *62,* 769-776.

Shan, L., Deng, K., Shroff, N.S., Durand, C.M., Rabi, S.A., Yang, H.C., Zhang, H., Margolick, J.B., Blankson, J.N., and Siliciano, R.F. (2012). Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. Immunity *36,* 491-501.

Shan, L., Yang, H.C., Rabi, S.A., Bravo, H.C., Shroff, N.S., Irizarry, R.A., Zhang, H., Margolick, J.B., Siliciano, J.D., and Siliciano, R.F. (2011). Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. J. Virol. *85,* 5384-5393.

Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J., and Siliciano, R.F. (2003). Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat. Med. *9,* 727-728.

Siliciano, J.D., and Siliciano, R.F. (2005). Enhanced culture assay for detection and quantitation of latently infected, resting CD4+ T-cells carrying replication-competent virus in HIV-1-infected individuals. Methods Mol. Biol. *304,* 3-15.

Simon-Loriere, E., and Holmes, E.C. (2011). Why do RNA viruses recombine? Nat. Rev. Microbiol. *9,* 617-626.

Singh, A., Razooky, B., Cox, C.D., Simpson, M.L., and Weinberger, L.S. (2010). Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. Biophys. J. *98,* L32-4.

Spivak, A.M., Salgado, M., Rabi, S.A., O'Connell, K.A., and Blankson, J.N. (2011). Circulating monocytes are not a major reservoir of HIV-1 in elite suppressors. J. Virol. *85,* 10399-10403.

Strain, M.C., Gunthard, H.F., Havlir, D.V., Ignacio, C.C., Smith, D.M., Leigh-Brown, A.J., Macaranas, T.R., Lam, R.Y., Daly, O.A., Fischer, M.*, et al.* (2003). Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence. Proc. Natl. Acad. Sci. U. S. A. *100,* 4819-4824.

Tamaru, H. (2010). Confining euchromatin/heterochromatin territory: jumonji crosses the line. Genes Dev. *24,* 1465-1478.

Temin, H.M. (1993). Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. Proc. Natl. Acad. Sci. U. S. A. *90,* 6900-6903.

Tyagi, M., Pearson, R.J., and Karn, J. (2010). Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction. J. Virol. *84,* 6425-6437.

Van Lint, C., Emiliani, S., Ott, M., and Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. EMBO J. *15,* 1112-1120.

Verdin, E., Paras, P.,Jr, and Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. EMBO J. *12,* 3249-3259.

Vicenzi, E., Dimitrov, D.S., Engelman, A., Migone, T.S., Purcell, D.F., Leonard, J., Englund, G., and Martin, M.A. (1994). An integration-defective U5 deletion mutant of human immunodeficiency virus type 1 reverts by eliminating additional long terminal repeat sequences. J. Virol. *68,* 7879-7890.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P., and Schaffer, D.V. (2005). Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. Cell *122,* 169-182.

Weinberger, L.S., Dar, R.D., and Simpson, M.L. (2008). Transient-mediated fate determination in a transcriptional circuit of HIV. Nat. Genet. *40,* 466-470.

West, M.J., Lowe, A.D., and Karn, J. (2001). Activation of human immunodeficiency virus transcription in T cells revisited: NF-kappaB p65 stimulates transcriptional elongation. J. Virol. *75,* 8524-8537.

Williams, S.A., Chen, L.F., Kwon, H., Ruiz-Jarabo, C.M., Verdin, E., and Greene, W.C. (2006). NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. EMBO J. *25,* 139-149.

Williams, S.A., and Greene, W.C. (2007). Regulation of HIV-1 latency by T-cell activation. Cytokine *39,* 63-74.

Wong, J.K., Hezareh, M., Gunthard, H.F., Havlir, D.V., Ignacio, C.C., Spina, C.A., and Richman, D.D. (1997). Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. Science *278,* 1291-1295.

Yang, H.C., Shen, L., Siliciano, R.F., and Pomerantz, J.L. (2009a). Isolation of a cellular factor that can reactivate latent HIV-1 without T cell activation. Proc. Natl. Acad. Sci. U. S. A. *106,* 6321-6326.

Yang, H.C., Xing, S., Shan, L., O'Connell, K., Dinoso, J., Shen, A., Zhou, Y., Shrum, C.K., Han, Y., Liu, J.O*., et al.* (2009b). Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation. J. Clin. Invest. *119,* 3473-3486.

Yu, Q., Konig, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J.M., and Landau, N.R. (2004). Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat. Struct. Mol. Biol. *11,* 435-442.

Ya-Chi Ho Last revised: September 2013

EDUCATION

Ph.D.	2013	Cellular and Molecular Medicine Program Johns Hopkins University School of Medicine Mentor: Robert F. Siliciano, M.D., Ph.D.
M.M.S	2007	Graduate Institute of Clinical Medicine National Taiwan University School of Medicine Mentors: Shan-Chwen Chang, M.D., Ph.D. Wei-Kung Wang, M.D., Sc.D.
M.D.	2002	National Cheng Kung University School of Medicine

PROFESSIONAL EXPERIENCE

Teaching assistant	2009	Pollard Scholar of Fundamentals of Genetics Cellular and Molecular Medicine PhD Program Johns Hopkins School of Medicine
Attending physician	2007-2008	Division of Infectious Diseases Department of Internal Medicine National Taiwan University Hospital
Clinical fellow	2005-2007	Division of Infectious Diseases Department of Internal Medicine National Taiwan University Hospital
Resident	2002-2005	Department of Internal Medicine National Taiwan University Hospital
Clinical rotation	2001	Visiting Medical Student Clerkship in Pulmonary Medicine, Consultative Cardiology Emergency Surgical Care and Neurology Duke University Medical Center
Clinical rotation	2000	Visiting Medical Student Clerkship in Medical Oncology Yale-New Haven Hospital
Research rotation	1999	Summer Research Student Laboratory of Lien-I Hor, Ph.D. Department of Microbiology and Immunology National Cheng Kung University School of Medicine American Bureau for Medical Advancement in China and Schering Plough Summer Research Grant

HONORS AND AWARDS

2013	First prize, Graduate Student Poster Competition Johns Hopkins University School of Medicine
2013	Young Investigator Award The 20th Conference on Retroviruses and Opportunistic Infections (CROI)
2011	Howard Hughes Medical Institute International Student Research Fellowship
2009	American Association for the Advancement of Science (AAAS)/Science Program for Excellence in Science
2008	Scholarship for clinicians doing research National Health Research Institute, Taiwan
2007	Scholarship for PhD Studentship Abroad Ministry of Education, Taiwan
2006	Best Teaching Resident Award National Taiwan University Hospital
2003-2005	Best Resident Award, Department of Internal Medicine National Taiwan University Hospital
2002	Dean's Award on graduation from National Cheng Kung University School of Medicine
2002	Best Intern Award National Cheng Kung University Hospital
2002	Doctor Tsong-Ming Tu's award
2002	*Phi Tau Phi* Scholastic Honor Society Award
2001	Dr. Andrew T. Huang Medical Education Promotion Grant
1999-2001	Presidential Award, National Cheng Kung University
1999	American Bureau for Medical Advancement in China and Schering Plough Summer Research Grant
1998	Tzu-Chi Charity Foundation Medical Scholarship
1996	Presidential Award, National Cheng Kung University

RESEARCH GRANT PARTICIPATION

Research fellow	2011	Analysis of non-induced HIV-1 provirus in latent reservoirs: a critical but neglected threat to HIV-1 eradication Howard Hughes Medical Institute International Student Research Fellowship
Junior Co-Principle Investigator	2013	Characterization of non-inducible proviruses Principle Investigator: Robert F. Siliciano CFAR-MDC NIH supplement working group National Institute of Health
Co-Principle Investigator	2006	Clinical manifestations and treatment result of patients with severe influenza infection in Taiwan Principle Investigator: Shan-Chwen Chang Centers for Disease Control, Taiwan

Assistant Investigator	2006	A clinical and epidemiological study of type A influenza in Taiwan Principle Investigator: Shan-Chwen Chang National Science Council, Taiwan

PUBLICATIONS

Peer reviewed articles

Ho Y.C., Liang S., Hosmane N.H., Wang J., Laskey S.B., Rosenbloom D. I. S., Lai J., Blankson J.N., Siliciano R.F. Replication-competent non-induced proviruses in the latent reservoir increase barrier to HIV-1 cure. **Cell.** In press (2013).

Yukl S.A., Busch M., Chun T.W., Douek D., Bentsen C., Eisele E., Haase A., **Ho Y.C**., Hütter G., Lee T.H., Justement J.S., Keating S., Li P., Murray D., Palmer S., Pilcher C., Pillai S., Price R.W., Rothenberger M., Schacker T., Siliciano J.D., Siliciano R.F., Sinclair E., Strain M., Wong J., Richman D.D., Deeks S.G. Challenges in detecting HIV persistence during potentially curative interventions: a study of the Berlin patient. **PLoS Pathog.** 9(5):e1003347 (2013).

Thayil S.M., **Ho Y.C.**, Bollinger R.C., Blankson J.N., Siliciano R.F., Karakousis P.C., Page K.R. *Mycobacterium tuberculosis* complex enhances susceptibility of CD4 T cells to HIV through a TLR2-mediated pathway. **PLoS One.** 7(7):e41093 (2012).

Wu U.I., Wang J.**T., Ho Y.C.**, Pan S.C., Chen Y.C., Chang S.C. Factors associated with development of complications among adults with influenza: A 3-year prospective analysis. **J. Formos. Med. Assoc.** 111(7):364-9 (2012).

Ho Y.C., Wang J.L., Wang J.T., Wu U.I., Chang C.W., Wu H.S., Chen C.H., Chuang Y.M., Chang S.C. Prognostic factors for fatal adult influenza pneumonia. **J. Infect.** 58(6):439-45 (2009).

Ho Y.C., Chang S.C., Lin S.R., Wang W.K. High levels of mecA DNA detected by a quantitative real-time PCR assay are associated with mortality in patients with methicillin-resistant *Staphylococcus aureus* bacteremia. **J. Clin. Microbiol.** 47(5):1443-51 (2009).

Sun H.Y., Kung H.C., **Ho Y.C.**, Chien Y.F., Chen M.Y., Sheng W.H., Hsieh S.M., Wu C.H., Liu W.C., Hung C.C., Chang S.C. Seroprevalence of hepatitis A virus infection in persons with HIV infection in Taiwan: implications for hepatitis A vaccination. **Int. J. Infect. Dis.** 13(5):e199-205 (2009).

Ho Y.C., Sun H.Y., Chen M.Y., Hsieh S.M., Sheng W.H., Chang S.C.. Clinical presentations and outcome of toxoplasmic encephalitis in patients with human immunodeficiency virus type one infection. **J. Microbiol. Immunol. Infect..** 41(5):386-92 (2008).

Ho Y.C., Shih T.F., Lin Y.H., Hiao C.F., Chen M.Y., Hsieh S.M., Sheng W.H., Hung C.C., Chang S.C. Osteonecrosis in patients with human immunodeficiency virus type one infection in Taiwan. **Jpn. J. Infecti. Dis.** 60(6):382-6 (2007).

Chuang Y.M., Tseng S.P., Teng L.J., **Ho Y.C.**, Hsueh P.R. Emergence of cefotaxime resistance in *Citrobacter freundii* causing necrotizing fasciitis and osteomyelitis. **J. Infect.** 53(3): e161-3 (2006).

Fang J.J., Shao C.P., **Ho Y.C.** and Hor L.I. Isolation and characterization of a *Vibrio vulnificus* mutant deficient in both extracellular metalloprotease and cytolysin. **Infect. Immun.** 69(9): 5943-8 (2001).

Oral Abstracts and Posters

Ho Y.C., Shan L., Wang J., Hosmane N.N., Blankson J.N., Silciano R.F. Replication-competent non-induced HIV-1 proviruses in the latent reservoir increase the barrier to HIV-1 cure. Johns Hopkins Center for AIDS Research (CFAR) Annual Meeting, Baltimore, United States. Poster presentation (2013).

Ho Y.C., Shan L., Wang J., Hosmane N.N., Blankson J.N., Silciano R.F. Characterization of non-induced HIV-1 proviruses dampens the hope for HIV-1 eradication. The 20th Conferences on Retroviruses and Opportunistic Infections. Atlanta, United States. Oral presentation (2013).

Ho Y.C., Blankson J.N., Siliciano R.F. Transwell limiting dilution co-culture as a novel and effective method for HIV-1 latent reservoir analysis. The 1114th General Meeting of the American Society of Microbiology, San Francisco, United States. Poster presentation (2012).

Yukl S., Chun T.W., Strain M., Siliciano J.D., Eisele E., Buckeit R., **Ho Y.C.**, Wong J., Busch M., Hutter G., Richman D.D., Siliciano R.F., Deeks S.G. Challenges Inherent in Detecting HIV Persistence During Potentially Curative Interventions. International Workshop on HIV & Hepatitis Virus Drug Resistance and Curative Strategies. Melia Sitges, Spain. Poster presentation (2012).

Ho Y.C., Siliciano R.F. Suberoylanilide hydroxamic acid (SAHA) does not reactivate additional HIV-1 proviruses from latent reservoirs in addition to those reactivated by phytohemagglutinin (PHA). Keystone symposia, Frontiers in HIV Pathogenesis, Therapy and Eradication (X8), Whistler, Canada. Poster presentation (2012).

Ho Y.C., Chang S.C. Prognostic factors for fatal pediatric influenza pneumonia. The 47th Annual Meeting of the Infectious Disease of America, Philadelphia, United States. Poster presentation (2009).

Ho Y.C., Wang W.K., Chang S.C. Quantification of methicillin-resistant*Staphylococcus aureus* in blood by a real-time polymerase chain reaction assay and its application in monitoring the treatment of methicillin-resistant *Staphylococcus aureus* bacteremia. Infectious Disease Society of Taiwan Annual Meeting, Taipei, Taiwan, oral presentation. The 108th General Meeting of American Society for Microbiology, Toronto, Canada, poster presentation (2007).

Ho Y.C., Wang J.T., Hung C.C., Chang S.C. Clinical profile of pathologically-proven toxoplasma lymphadenitis in non-HIV infected patients. Infectious Disease Society of Taiwan Annual Meeting. Oral presentation. (2006).

Ho Y.C., Hor L.I. Roles of protease and cytolysin in tissue damage caused by *Vibrio vulnificus*. American Bureau for Medical Advancement in China and Schering-Plough Summer Research Scholarship Conference, Taipei, Taiwan. Oral presentation (1999).